THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

Section B

Vol. 63, Part 6 — 1 June 1950 — No. 366 B

The Behaviour of the Cathode Spot on an Undisturbed Liquid Surface of Low Work Function

By K. D. FROOME *
Department of Physics, Imperial College, London

MS. received 19th September 1949

ABSTRACT. The cathode spot of transient arcs using the liquid sodium–potassium alloy for cathode has been studied by means of the Kerr cell apparatus capable of taking a sequence of photographs. It is found that for a given current the spot takes, as in the case of mercury, the form of a line or broken line of total length roughly proportional to the current. This is also true for a spot in a magnetic field. Again as for mercury, if the current rises slowly the length of line increases proportionally to the current. If the rate of growth of current is greater than a value lying approximately between 10^8 and 2×10^8 amp/sec., then fresh spots form and spread out radially from the point of formation into thin semicircular or circular lines moving with a maximum radial velocity of about 10^4 cm/sec. The apparent current density of emission lies between 2·5 and 5×10^6 amp/cm^2, and consequently is higher than for mercury.

§ 1. INTRODUCTION

The results described in this paper have been taken by means of the Kerr cell apparatus capable of taking a series of superimposed photographs which show a number of positions of the cathode spot as it moves over the cathode. The interval between each exposure can be varied from a fraction of a microsecond to many microseconds, and the number of exposures of the cathode spot taken during a discharge can be adjusted from one to as many as are required to build up a complete picture of the behaviour of the emitting area. The exposure can be varied independently of either interval or number. This apparatus has been described elsewhere (Froome 1948).

Observations are made by means of a low-power microscope arranged to look down on to the liquid cathode surface, the Kerr cell itself being placed just outside the microscope eyepiece. The actual optical arrangement used is shown in Figure 1 of a paper describing details for an undisturbed mercury cathode (Froome 1949); reference should be made to this paper for other allusions to mercury cathodes.

The transient discharges studied are obtained by discharging a single condenser, or a number of condensers, in the form of artificial lines, charged to approximately 100 volts through the tube and series impedance. The former is used when varying currents are required, and the artificial lines can be used to produce constant current pulses of 10 to 450 amperes.

---

* Now at the National Physical Laboratory, Teddington, Middlesex.

The discharge is termed an arc because the cathode spot emits a large number of electrons, and because there is a low overall potential drop. The term cathode spot is used to denote the whole emitting area on the cathode irrespective of its exact form.

The liquid alloy used has a constitution of approximately 25% sodium and 75% potassium, but almost identical results can be obtained with any other proportion provided it is liquid at room temperature. The alloy has been chosen for two reasons. Firstly, it is liquid at room temperature, and a liquid surface is desirable because it is reproducible; secondly, it has a low work function, whereas mercury has a high work function (4·5 volts). The curves given by Thomson and Thomson (1928) for photoelectric emission from the alloy indicate a work function of about 2 volts.

In order to use this alloy, the discharge tube must be sealed off. It was not found possible to purify the alloy by the technique of evaporation usual for the alkali metals, for this separated the constituents of the alloy and deposited potassium in an irremovable film over the discharge tube window. The following technique eventually proved entirely satisfactory.

The tube is made with a long thin glass pumping line attached. This line consists of a number of bulbs of about one inch diameter separated from each other by a narrow capillary neck. The end farthest from the discharge tube (and nearest the liquid air trap) is fitted with a cone joint device by means of which the previously prepared and consequently rather dirty alloy can be introduced into the vacuum line. The whole is out-gassed as thoroughly as possible and filled with argon to the operating pressure (0·1 to 0·01 mm.). Then the liquid alloy is tipped into the bulb farthest from the discharge tube and the whole sealed off from the pumps. The alloy is then passed from one bulb to the next through the narrow joining capillary tubes, and by the time it reaches the discharge tube it appears brilliantly clean. The discharge tube is then sealed off from the cumbersome remaining line of bulbs.

The type of discharge tube used is almost the same as that described earlier (Froome 1949, Figure 1), but no amalgamated copper foil is wrapped round the inside of the cathode neck. Consequently the alloy surface is always somewhat curved, and for this reason some of the photographs shown are in focus only in one or two successive exposure positions.

Finally, the cathode is cleaned further by running the tube for several hundred flashes at high peak current. This is absolutely essential if reproducible results are to be obtained. The cathode spot has a strong cleansing action upon the cathode.

In all the Kerr cell photographs reproduced the cathode spot has started from the bottom of the picture. Hence each photograph shows a number of positions of the emitting cathode area separated by the time interval stated, with time increasing as the cathode spot moves upwards and, in the case of rapidly rising currents, outwards as well. The discharge is initiated by means of a brief high tension pulse applied outside the discharge tube near the cathode. It is found that the cathode spot almost invariably forms on the liquid surface adjacent the glass wall of the tube nearest the position of application of the triggering pulse. In some of the photographs (notably those shown in Plate II) the curved surface of the glass cathode wall can be seen in black outline along the bottom of the picture.

§2. RESULTS WITH CONSTANT CURRENT PULSES

It is difficult with this alloy to obtain results from an undisturbed surface. The emitting area rapidly breaks up into a large number of minute separate spots. These 'unit spots' seem to form in self-made dents in the surface. The same behaviour is observed for mercury, but takes considerably longer.

Figure 1 (Plate I) shows a photograph with exposure 1 microsecond taken 4 microseconds from the start of a constant current discharge of 410 amperes. It is seen that even in this short time unit spots have formed, whereas the corresponding time for mercury would have been 20–30 microseconds or more.

Such observations suffer from the limitations imposed by these dents (Froome 1949); and much of the work for constant current pulses has therefore been undertaken with the cathode spot in a strong magnetic field, since this produces a movement of the cathode spot sufficiently rapid to inhibit unit spot formation.

For the cathode spot in a magnetic field we again find, as for mercury, a marked reluctance to travel with a velocity of more than 10^4 cm/sec. Also, if the velocity is much below this value, excessive formation of unit spots takes place, and these are left behind the main emitting area until they eventually extinguish.

Nevertheless, there is a marked similarity in behaviour to that for mercury. Figure 2 (*a*)–(*d*) (Plate I) shows sequences of exposures of the cathode spot for a constant current of 90 amperes at two magnetic field strengths and with two different exposures. The magnetic field is across the picture, the spot moving in exactly the opposite direction to the ponderomotive force in the emitted electrons. In Figure 2 (*a*) the field is 900 gauss, in (*b*) it is 1,800 gauss. The exposure in each case is 0·1 microsecond, and the interval between exposures is 4·2 microseconds. At the higher field it is seen that the emitting area is in the form of a line similar to that for mercury, but even at this field strength, with the consequent high velocity of the spot, there is evidence of the formation of a few unit spots which have been left behind in the wake of the main emitting area. At the lower field the line has become extremely ragged, or broken, due to such spot formation, and thus it is difficult to assess the actual emitting area.

Figure 2 (*c*) shows a sequence, with longer exposures (0·3 microsecond) but with the same interval between each, of another discharge again in a field of 900 gauss. The raggedness is more clearly seen—as are the numbers of unit spots. Figure 2 (*d*) shows a spot in the higher field (1,800 gauss) with this longer exposure. The line-like structure of the main emitting area is again seen, but there is also evidence of unit spots which have broken off and been left behind. This behaviour is due to the dents such unit spots apparently make in the liquid surface. Once a dent has formed ***the spot is not capable of rapid movement***; this applies to mercury also.

Figure 3 (*a*)–(*c*) (Plate I) shows multiple exposures of cathode spots carrying various currents in a magnetic field of 1,300 gauss. The exposure in each is 0·3 microsecond, and the interval between consecutive exposures is 4·2 microseconds; (*a*) shows the spot for a current of 14 amperes, (*b*) for 60 amperes, and (*c*) for 165 amperes. The emitting area is seen to be chiefly a line for the two lower currents, but for the highest it consists of a very ragged (broken) line plus a large number of unit spots in its wake. It is also seen that these spots have a life-time of the order of the interval between consecutive exposures, namely about 4 microseconds.

Figure 4 shows graphically a plot of the current against length of this emitting line. As for mercury, this curve is obtained from a large number of different results. In this case, however, only those results which show the best line-structure have been included, and a correction has been attempted for the probable effect of the residual unit spots.

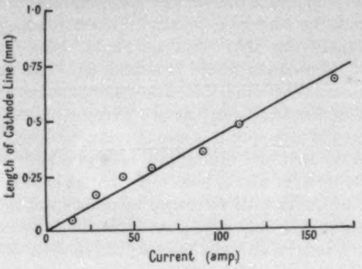


Figure 4.

As for mercury, we see that the length, *l*, of the emitting line is proportional to the current, *i* (amp.), where $l = Ki$ cm., and $K \simeq 4{\cdot}1 \times 10^{-4}$ cm/amp^{-1}. This is under one-half the value obtained for mercury.

§ 3. RESULTS WITH VARYING CURRENTS

The results with varying currents show a strong similarity to the corresponding behaviour of mercury. As before, these results are obtained with the spot not in a magnetic field, in order to maintain simple conditions.

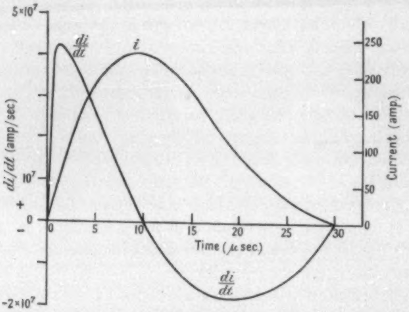


Figure 6.

Figure 5 (Plate II) shows the appearance of the cathode spot for the electrical conditions shown in Figure 6. It will be remembered that for mercury these conditions produced the formation of fresh spots. Figure 5 shows three exposures of 0·3 microsecond separated by 4·2 microseconds, the first being taken 2 microseconds from the initiation of the discharge. It is seen that only the first shows anything of a line-like structure, and that in the second and third slow-moving

PROC. PHYS. SOC. VOL. 63, PT. 6—B (K. D. FROOME)

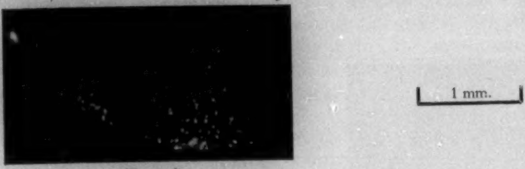


Figure 1.

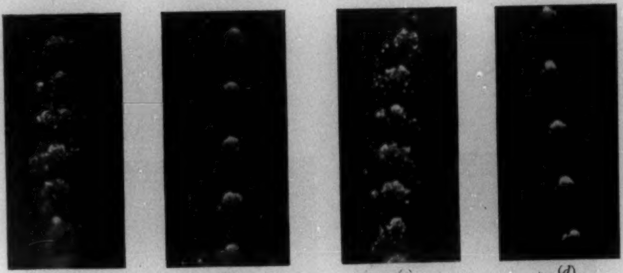

(a) (b) (c) (d)

Figure 2.

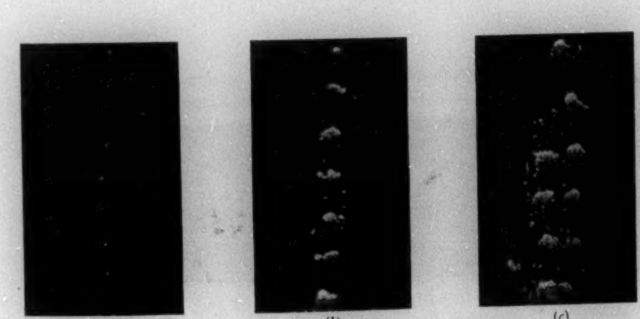

(a) (b) (c)

Figure 3.

PLATE I.

Figure 5.

1 mm.

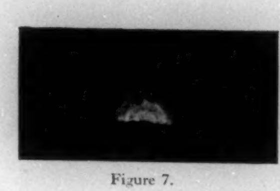


Figure 7.

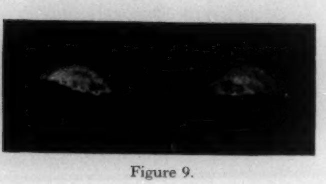
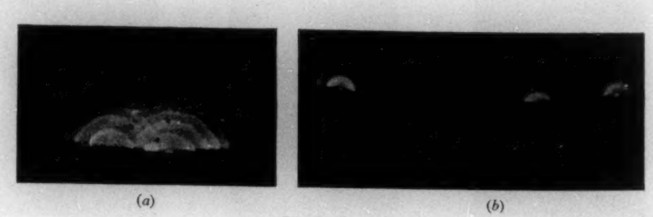


Figure 9.



(*a*) (*b*)

Figure 10.

PLATE II.

unit spots have formed; consequently in each of the last two exposures one can scarcely distinguish the positions of individual spots.

We have already seen that for a spot in a magnetic field the length of the emitting line for a given current is under one-half that of mercury. Hence, since the spot is reluctant to move much faster than 10^4 cm/sec., one can expect that at very high values of rate of growth of current the cathode spot will be forced to move in the form of a semicircular line at its maximum velocity, or to form new spots.

When the emitting area does take this form of a semicircular line moving outward with a maximum permitted velocity of 10^4 cm/sec., its maximum length L after t seconds will be given by $L = 10^4 \pi t$ cm., or, after t sec., the maximum current i emitted by such a line will be given by L/K where $K \simeq 4 \times 10^{-4}$. Hence the maximum current $i = 8 \times 10^7\, t$, and therefore $di/dt = 8 \times 10^7$ amp/sec. represents the maximum rate of growth of current which we can expect to be maintained by a single semicircular spot. (The spot is almost invariably semicircular since it starts adjacent to the wall of the glass tube and can therefore spread out only into the tube. If the spot does form on the cathode surface away from the tube wall, then it can form a completely circular line, so that the maximum rate of growth of current sustainable by such a spot will be twice the above value.)

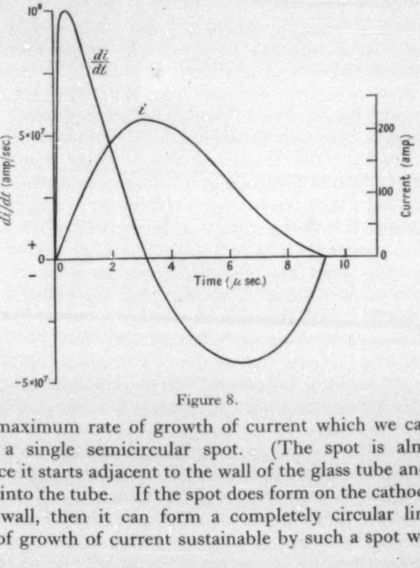


Figure 8.

Figure 7 (Plate II) shows a sequence of four exposures (exposure 0·1 microsecond, interval 1·0 microsecond, the first exposure occurring just under a microsecond from the start of the arc) of the cathode spot obtained by discharging a 20 μF. condenser initially at 60 volts through the tube in series with a total inductance of about 0·3 μH. with negligible total circuit resistance. The peak value of rate of growth of current is 10^8 amp/sec. The cathode spot has formed as a single spot adjacent to the tube wall at the bottom of the picture, and then disintegrated into a semicircular line moving radially outward with a velocity of about 10^4 cm/sec. The characteristics and time of the discharge are shown in Figure 8.

Figure 9 (Plate II) shows the result with the same exposure conditions, with the condenser charged to 108 volts. The peak value of rate of growth of current is now $2\cdot4 \times 10^8$ amp/sec. The actual values of current and rate of growth of current can be obtained from Figure 8 by multiplying by a factor 2·4. The time of the discharge is still as shown in Figure 8. In Figure 9 two spots have formed, disintegrating into semicircular lines moving with a velocity of nearly 10^4 cm/sec. Five positions are shown, the first appearing about 0·5 microsecond after the start of the discharge, so that just over 4 microseconds of discharge time are shown.

Figure 10 (*a*) and (*b*) (Plate II) shows two photographs of spots obtained with the same exposure and circuit conditions as before, but with the condenser charged to 144 volts. The time of the discharge is still as shown in Figure 8, and actual values of current and rate of growth of current can be obtained by multiplying the values here shown by a factor of about 3·4. In (*a*) two spots have simultaneously formed and spread out into lines moving with an initial velocity of about $1\cdot3 \times 10^4$ cm/sec. The cathode spot is shown for about the first five microseconds in about six positions—the first exposure of the right-hand spot occurring practically simultaneously with the formation of the spot, the left-hand spot having formed about 0·1–0·2 microsecond later.

In (*b*), three spots have formed simultaneously and spread, initially moving outwards at 10^4 cm/sec. The first few microseconds are shown in three or four exposures, the first exposure being about 0·1–0·2 microsecond after the start.

Hence we see that a spot can support a maximum rate of growth of current of about 10^8 amp/sec. (if it starts outwards from the wall of the tube). If higher rates of growth of current are required, then extra spots appear simultaneously. This is approximately what was predicted from the study of the spot in the magnetic field. If the spot should form away from the wall of the tube then the critical value would be 2×10^8 amp/sec.

The formation of fresh spots should not be confused with the appearance of large numbers of unit spots formed when the velocity of the emitting line decreases to slightly less than 10^4 cm/sec. These unit spots are not fresh spots, but broken-up pieces of emitting line. Fresh spots only form under the conditions of rapid growth of current. If the current is relatively slowly rising (as was the case in Figure 5) and unit spots have formed, the size of the emitting area increases by the breaking up of unit spots into greater numbers, and not by the spontaneous formation of fresh spots. If the current is relatively slowly rising, but the velocity of the emitting area is sufficiently high for unit spot formation not to take place (e.g. with the spot in a magnetic field), then the emitting area preserves its line-like structure, increasing proportionally with increasing current. This suggests that when unit spot formation has taken place each unit spot consists of a piece of line in the bottom of a dent, and that if the current is increasing the pieces of line in some of the dents increase, resulting in instability of the dents, so that they break into two or more separate dents. If the current is decreasing the total length of emitting line decreases proportionally, and the number of unit spots decreases.

Returning to the discussion of rapidly rising currents with the cathode spot not in a magnetic field, Figure 11 shows a plot of length of emitting line against current for a number of discharges similar to those shown in Figures 7 to 10

(Plate II). Again we find approximate linearity between the length of line, l, and current, i (amp.), such that $l = Ki$ cm., where $K = 3{\cdot}7 \times 10^{-4}$ cm/amp^{-1}.

In view of the experimental difficulties, this value of K may be taken as identical with that obtained for the spot in a magnetic field.

§ 4. DISCUSSION AND CONCLUSIONS

It is a curious feature that, whereas for mercury (Froome 1949) the fresh spots form in front of the previous ones at high values of rate of growth of current, with the liquid alloy they almost invariably form along the edge of the cathode adjacent to the glass wall. Further, there is less time lag between the achievement of the high value of rate of growth of current and the formation of the fresh spots required to sustain it. For mercury this time lag may be of the order of a microsecond, whereas for the liquid alloy it appears to be nearer 0·1 microsecond.

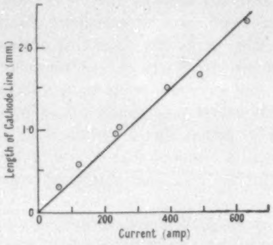


Figure 11.

Most of the conclusions given for mercury apply also to the liquid alloy; the most significant feature is that when the emitting area takes the form of a line, it is shorter than for mercury at the same current. The apparent width of the line for the alloy, as far as can be estimated, is between 10^{-3} and 5×10^{-4} cm., leading to an apparent emission current density of $2{\cdot}5 \times 10^{6}$ to 5×10^{6} amp/cm^{2}, whereas the apparent current density for mercury is only slightly in excess of 10^{6} amp/cm^{2}.

There is considerably less copious evolution of vapour from the alloy cathode than from the mercury cathode, and it may be that for mercury the large amount of vapour emitted renders the surface at the cathode spot indeterminate and lowers its effective work function.

For the liquid alloy the emitting line is easily broken up into unit spots even when moving at a velocity of 10^{4} cm/sec. For mercury the corresponding velocity is about 2×10^{3} cm/sec. It is well known that the cathode spot exerts a considerable net downward pressure on the cathode under it, so that for the less dense liquid alloy rupture into dents (and hence unit spots) would take place more rapidly.

All the photographs here reproduced are of a region of dense and highly excited vapour shot off from the cathode at the emitting area. It is not known how far out from the cathode this region extends, or if there is a 'dark space' between this region and the cathode, as is the case for mercury (see Smith 1946). For mercury this 'dark space' extends outwards for about 10^{-3} cm., and as the width of the emitting line appears also to be 10^{-3} cm. for mercury, the writer has suggested (Froome 1949) that the actual emitting line may be thinner than this and that a method of estimating its real size (as opposed to its apparent size) is to determine the maximum current which the spot will carry. For mercury the value is between 0·5 and 0·1 amp., and if at this value the spot is symmetrical, i.e. its length is equal to its breadth, the suggested real emission current density is 2×10^6 to 10^7 amp/cm². For the liquid alloy the minimum current lies between the same values; hence, by applying the same argument, the real emission current density may lie between the extraordinarily high values of $1{\cdot}2 \times 10^7$ and 6×10^7 amp/cm², if we assume that for a current of i amp. the length of the emitting line l is given by $l = Ki$ where $K \simeq 4{\cdot}0 \times 10^{-4}$ cm. amp^{-1}.

Even if we do not assume the real value of emission current density to be as high as that given above, and take only the apparent value of $2{\cdot}5$ to 5×10^6 amp/cm², we find that Langmuir's (1923) original theory of field emission is well substantiated without the assumption of an enhanced electric field at the cathode due to surface bumps. This is the first time that emission densities as high as those predicted by Langmuir's theory have been observed.

Langmuir suggested that since electrons have a high mobility, the positive ion current to the cathode, across the cathode fall of potential, must be positive space-charge-limited. This means that a certain field must exist in order that this current can pass, and he assumed that this field would be high enough to extract the electron current by 'cold' or 'high field' emission from the cathode surface.

Langmuir's well known space-charge formula reads

$$J_+ = \frac{1}{2\pi}\left(\frac{2\boldsymbol{e}}{M_+}\right)^{1/2}\frac{V^{3/2}}{d^2}$$

where J_+ is the positive ion current density, V the cathode fall of potential, d the thickness of the cathode dark space, M_+ mass of the positive ions, and $\boldsymbol{e}$ the electronic charge.

To pass this current a field X is needed where

$$X = 4\pi\left(\frac{M_+}{2\boldsymbol{e}}\right)^{1/4} J_+^{1/2} V^{1/4} = (3{\cdot}28 \times 10^7)^{1/2} M^{1/4} V^{1/4} \text{ volt/cm.};$$

M is now the atomic weight of the positive ions, and J_+ and V are in practical units.

For the liquid alloy this reduces to $X = 2{\cdot}6 \times 10^4 J_+^{1/2}$, assuming for convenience a cathode fall of 10 volts.

To obtain the electron emission density produced by this electric field, the equation of Fowler and Nordheim (1928) is best used. This states that the emission current density J_- for an electric field X is given by

$$J_- = \frac{6{\cdot}2 \times 10^{-6} X^2}{\phi} \exp\left(\frac{-6{\cdot}8 \times 10^7 \phi^{3/2}}{X}\right) \text{ amp/cm}^2$$

where ϕ, the work function, is 2 volts.

Figure 12 shows the electron emission density plotted against various incoming positive ion densities. It is seen that as soon as the positive ion current density attains a value of 10^6 amp/cm² an electron emission density greater than this value is obtained. The results described show apparent emission densities of up to 5×10^6 amp/cm². At such a value a positive ion current density of 10^6 amp/cm² is quite conceivable, so that the results described may be taken to verify Langmuir's theory.

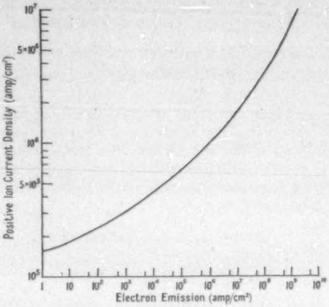


Figure 12.

Further, as for mercury, there is an apparent afterglow, lasting a microsecond or so, in the wake of a moving cathode spot.

ACKNOWLEDGMENTS

The writer wishes to express his thanks to Professor Sir George P. Thomson for many valuable discussions, and to Mrs. K. D. Froome for preparation of the diagrams.

REFERENCES

Fowler, R. H., and Nordheim, T., 1928, *Proc. Roy. Soc.* A, **118,** 229.
Froome, K. D., 1948, *J. Sci. Instrum.*, **25,** 371; 1949, *Proc. Phys. Soc.* B, **62,** 805.
Langmuir, I., 1923, *Gen. Elect. Rev.*, **26,** 731.
Smith, C. G., 1946, *Phys. Rev.*, **69,** 96.
Thomson, J. J., and Thomson, G. P., 1928, *Conduction of Electricity through Gases*, Vol. I (Cambridge: University Press), p. 479.

The Calculation of Magnetic Lens Fields by Relaxation Methods *

By M. B. HESSE†
Imperial College, University of London

Communicated by G. I. Finch; MS. received 6th July 1949, *and in amended form 26th October* 1949

ABSTRACT. The field and lens constants are calculated for two typical magnetic lenses as used in electron microscopes, using the relaxation method developed by Southwell for the solution of potential problems. The results are compared with those based on simple analytic approximations to the field distribution in the work of Glaser and Ramberg, and are shown to agree closely as regards those characteristics on which the discussion of the performance of the lens is based.

---

§ 1. INTRODUCTION

It is an advantage in many problems connected with the design of electron lenses to have an accurate knowledge of the field distribution in the neighbourhood of the pole pieces. The determination of the characteristics of a lens, including the aberrations, can be carried out if the field distribution along the axis of symmetry of the lens is known. Bertram (1940, 1942) has solved the field equation analytically with boundary conditions approximating to those of a practical lens with unsaturated pole pieces, and finds a field expression of the form

$$H = H_{\mathrm{m}} \operatorname{sech}^2 az.$$

Glaser (1941) has exhaustively studied the trajectories due to a field of the type

$$H = \frac{H_{\mathrm{m}}}{1+(z/a)^2}.$$

No attempt has, however, been made to solve theoretically the problem of the magnetic lens with saturated pole pieces, although this is the condition under which lenses are worked in practice.

This paper describes the solution of the field equations by the relaxation method developed by Southwell and his collaborators (Southwell 1946) for the numerical solution of potential problems. The calculation has been carried out for both unsaturated and saturated pole pieces.

§ 2. EQUATIONS FOR THE FLUX FUNCTION

In the first part of this paper it will be assumed that the relation between the magnetic field **B** and the magnetizing force **H** is linear, that is, $\mathbf{B} = \mu\mathbf{H}$, where μ is the constant magnetic permeability. This is only true over a restricted range of **H** for materials used in practice, and the modifications required when μ is not constant will be discussed below.

The application of the relaxation method to a problem of field distribution is simplest when the field equations are expressed in a form containing a function ψ

---

* This work formed part of a thesis submitted to the University of London for the Ph.D. degree.
† Now at Royal Holloway College, University of London.

which is constant along the flux lines. ψ is defined by the equations

$$\mu H_z = -\frac{1}{r}\frac{\partial\psi}{\partial r}, \qquad \mu H_r = \frac{1}{r}\frac{\partial\psi}{\partial z}. \qquad \ldots\ldots(1)$$

The equation for ψ is

$$\frac{\partial^2\psi}{\partial r^2} + \frac{\partial^2\psi}{\partial z^2} - \frac{1}{r}\frac{\partial\psi}{\partial r} = \frac{0{\cdot}4\pi nI}{\Delta} r, \qquad \ldots\ldots(2)$$

where nI is the number of ampere turns in the coil and Δ is the total cross-sectional area of the windings.

Physically, the value of ψ at any point gives the magnetic flux passing through a circle which lies in a plane perpendicular to the axis, with its centre on the axis and its circumference passing through the given point.

At a boundary between media with different values of μ the tangential component of **H** is continuous, and the normal component of **B** is continuous.

Thus at a point on the boundary

$$(H_t)_1 = (H_t)_2; \qquad (\mu H_n)_1 = (\mu H_n)_2. \qquad \ldots\ldots(3)$$

From equations (1)

$$(\psi)_1 = (\psi)_2; \qquad \left(\frac{1}{\mu}\frac{\partial\psi}{\partial n}\right)_1 = \left(\frac{1}{\mu}\frac{\partial\psi}{\partial n}\right)_2. \qquad \ldots\ldots(4)$$

§ 3. THE RELAXATION METHOD

The solution of a potential distribution is given by the relaxation method in terms of a function defined by computed values, which satisfies the finite difference equation corresponding to the field at a finite number of points. The number may be increased indefinitely to obtain the degree of accuracy required, the only limitation being the labour involved in solving the system of simultaneous equations. The relaxation method is essentially a device for solving a large number of such equations quickly, and in such a way that mistakes can be detected and rectified at any stage.

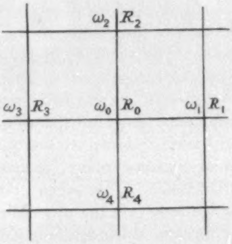


Figure 1.

The points at which the values of the function are to be calculated are most conveniently taken at the corners of a square mesh, as shown in Figure 1. The value of the function ω_0 at any general point of the field may be found in terms of the values ω_1, ω_2, ω_3, ω_4, at surrounding points, and the mesh length h. Special equations are needed at boundary points between media of different permeability.

We require to solve the differential equation for the flux function over the field of the lens. As the function will be computed in arbitrary units, equation (2) is replaced by

$$\frac{\partial^2\omega}{\partial r^2}+\frac{\partial^2\omega}{\partial z^2}-\frac{1}{r}\frac{\partial\omega}{\partial r}+rZ=0. \qquad \ldots\ldots (5)$$

ψ is then given by

$$\frac{\psi}{\omega}=\frac{0{\cdot}4\pi nI/\Delta}{Z}, \qquad \ldots\ldots (6)$$

where Z is a constant whose value depends on the units in which ω is calculated.

At ordinary points of the field the first-order finite difference equation corresponding to (5) is

$$\omega_1+\left(1-\frac{h}{2r}\right)\omega_2+\omega_3+\left(1+\frac{h}{2r}\right)\omega_4-4\omega_0+h^2Z_0r=0.$$

This is obtained by expanding ω in a Taylor series at the point O and neglecting terms containing powers of h equal to or higher than the third. It is therefore accurate to the square of the mesh length.

The first step in the process of solution is to guess a distribution of ω throughout the field. Then the residual

$$R_0=\omega_1+\left(1-\frac{h}{2r}\right)\omega_2+\omega_3+\left(1+\frac{h}{2r}\right)\omega_4-4\omega_0+h^2Z_0r \qquad \ldots\ldots (7)$$

is calculated for each point of the mesh. The residuals are now reduced to zero by subtracting $R/4$ from ω at the point where R is greatest, and modifying the surrounding points accordingly. The process is repeated until all the residuals lie between -2 and $+2$ (where the units of this residual are those of the last digit or decimal of ω), after which no change in the values of ω will reduce them further. Values of ω which are fixed by the boundary conditions are of course never changed. Finally R is calculated from equation (7) for each point again, and any mistakes are detected by the occurrence of residuals greater than 2, and can be rectified by reducing these as before. At the final stage the signs of the residuals should be distributed at random over the field, and the sum of the residuals should be almost zero. This reduces the possibility of cumulative errors in any part of the field. The accuracy of the solution in any region can be increased by reducing the mesh length, or by taking account of higher order terms in the Taylor series when forming the difference equations.

Special difference equations are required for points near a boundary where the distances between neighbouring points are unequal, the so-called 'irregular stars'. Also it is convenient at some points to use a square net with sides inclined at an angle of 45° to the z axis. This is required near diagonal boundaries and in regions where a change-over to a net of smaller mesh takes place, and is also useful for interpolating the value at the centre of a coarse mesh. At boundary points it is assumed as a first approximation that the value of μ was everywhere of the order of 1,000, so that $1/\mu$ can be neglected in comparison with unity, giving a possible error of 0·1% in the value of ω. Then the boundary equation (4) reduces to $(\partial\omega/\partial n)_{\text{air}}=0$. The region outside a sharp corner presents some difficulty. The flux lines tend to concentrate round the corner, and the value of the flux function varies rapidly so that the Taylor series converges slowly. If the corner were mathematically sharp the flux density would be infinite there, but in fact the radius of curvature at the corner of the iron pole pieces is finite, and the mesh in this region should be at least as small as the radius.

The difference equations for all these special cases are given in the Appendix.

§4. CALCULATION OF THE LENS FIELDS AND TRAJECTORIES

The flux distribution was investigated for a simple lens such as might be used for the objective of an electron microscope, the permeability being assumed infinite. The boundary conditions of the problem are given by the results of the previous section, and the value of ω along the axis is given by

$$\frac{\partial\psi}{\partial r} = -\mu r H_z, \qquad \frac{\partial\psi}{\partial z} = \mu r H_r.$$

ψ and ω are therefore zero on the axis.

At the first stage of the calculation the value of Z was not defined, but a reasonable distribution of ω was assumed, and at each mesh point the function

$$\omega_1 + \left(1 - \frac{h}{2r}\right)\omega_2 + \omega_3 + \left(1 + \frac{h}{2r}\right)\omega_4 - 4\omega_0,$$

or the equivalent expression for boundary points, was calculated. Then the average value of this residual was obtained and Z chosen so that the total residual over the area of the windings was zero. This device was found to lighten the work considerably, since the only fixed values of ω at which an accumulated residual could be 'thrown away' lie on the axis at one corner of the field, opposite the air gap (marked X in Figure 2).

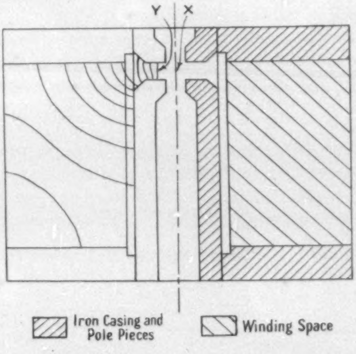


Figure 2. Section across objective lens I showing flux lines in the windings.

Rigorous justification of the method of solution is not necessary; the method is immaterial if the final values of ω satisfy the original set of linear equations, which have a unique solution. The error involved in approximating to the differential equation by finite difference equations, and the possible accumulation of errors in the relaxation process itself, however, has to be discussed. It was mentioned above that the finite difference equations are an accurate representation of the differential equation if powers of the mesh length higher than the third can be neglected. This can be tested in the case of any solution by halving the mesh length and finding the values of ω at the new nodes. If the values previously obtained in the larger mesh remain unaltered, the neglect of terms in the Taylor series is justified; if not, the mesh length must be further reduced until two successive calculations correspond.

In unsaturated iron the permeability coefficient is about 1,000. Thus three-figure accuracy can be obtained in the values of ω if infinite permeability is assumed. This accuracy can be maintained over the whole field irrespective of the absolute value of ω, as may be shown by considering two sets of solutions,

$$u_1, u_2, u_3, \dots u_n,$$

and

$$u_1+\delta u_1, u_2+\delta u_2, u_3+\delta u_3, \dots u_n+\delta u_n,$$

where $\delta u_r/u_r<\epsilon$ and ϵ is the possible proportional error of u_r in one region of the field. The only fixed boundary values of ω are those on the axis, where ω is zero, and since the field represents a set of simultaneous linear equations in $\omega_1, \omega_2, \dots \omega_n$, any set of solutions may be multiplied in the same ratio to give another set of solutions. Thus

$$\frac{\delta u_1}{u_1}=\frac{\delta u_2}{u_2}=\dots=\frac{\delta u_n}{u_n},$$

and the possible percentage error remains constant over the whole field.

The absolute value of ω falls rapidly towards zero near the axis, so that the values of ω must be calculated to a greater number of decimal places near the axis than is necessary near the pole pieces (see Table 3), but since the possible percentage error remains constant, three-figure accuracy can be obtained throughout the field. Another possible source of error arises from the fact that the only fixed values of ω lie on the axis, and errors in the rest of the field are likely to be cumulative. The sum of the residuals in any arbitrary area of the distribution should be numerically less than 10% of the number of mesh points in order to reduce this possibility.

The absolute value of H_z along the axis may be found in terms of the number of ampere turns in the coil by an application of Ampère's law

$$\int_C \mathbf{H}\,.\,\mathbf{ds}=0{\cdot}4\pi nI,$$

where C is a closed curve lying along the axis and round the outside of the iron case where the field is zero. Using equations (1) and expanding ω in a Taylor series we obtain

$$(h^2H_z)_{r=0}=2K\omega(h) \qquad \dots\dots(8)$$

and

$$H_m\int_{-\infty}^{\infty}\frac{\omega(h)}{\omega_m}dz=0{\cdot}4nI; \qquad \psi=K\omega,$$

where ω_m is the value of $\omega(h)$ at $r=h$, $z=0$, and H_m is the value of H at $r=0$, $z=0$.

The distribution of H_z for lens I is shown in Figure 3.

The calculation for this lens indicated that one of the flux lines (Y in Figure 2) is straight and parallel to the axis at the centre of the pole-piece gap. If this is assumed to be generally true for all lenses with similarly shaped pole pieces within the limits of accuracy required, the work of calculating the field distribution is considerably reduced. The distributions of H_z were found by this method for similar lenses in which the ratio of the gap to the inner diameter d of the pole pieces varies; these are shown in Figure 3, together with the distributions assumed by Glaser (1941) and Ramberg (1942).

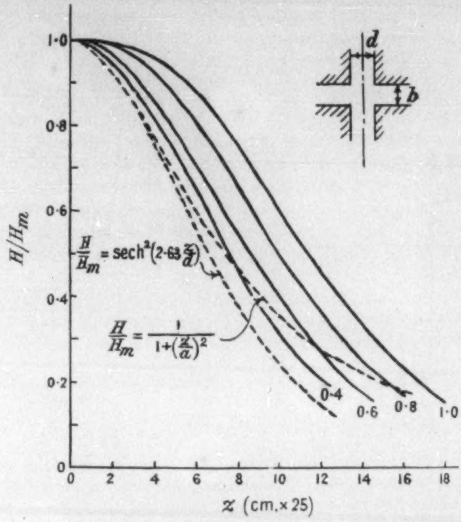


Figure 3. Variations of the field distribution with the pole piece ratio b/d.

The equation for the trajectories of paraxial electrons in a magnetic lens is

$$\frac{d^2r}{dz^2} = \frac{-e}{8mc^2V} rH^2. \qquad \ldots\ldots (9)$$

A numerical step-by-step method of solution is suitable, as the values of H are available at intervals along the axis. Neglecting terms of order h^4, equation (9) becomes

$$r_{z-2h} + r_z - 2r_{z-h} = h^2 \frac{e}{8mc^2V} r_{z-h} H^2_{z-h}.$$

Putting $$H_{z-h} = H_{\mathrm{m}} \frac{\omega(h)}{\omega_{\mathrm{m}}} \qquad \text{and} \qquad \frac{eH_{\mathrm{m}}}{8mc^2V} = k,$$

we obtain $$r_z = \left\{2 - kh^2 \frac{\omega^2(h)}{{\omega_{\mathrm{m}}}^2}\right\} r_{z-h} - r_{z-2h}.$$

From this equation the radial distance of an electron is calculated at intervals $z = h$. $\omega(h)/\omega_{\mathrm{m}}$ is accurate to three figures, and five figures may be retained in the values of r_z. Since the number of steps is never greater than fifty, the final values of r_z may safely be regarded as accurate to three figures.

To find the focal length, the initial path of the electron before it enters the field is assumed to be at a constant distance from the axis. The focal length is defined by analogy with geometrical optics. It is obtained by producing the final line of the path backwards to cut the initial path (Figure 4). s is the apparent shift of the centre of the lens system and f and s must be used in all calculations involving the formulae of geometrical optics. These formulae will only be valid if the object distance is greater than s.

Figure 5 shows the focal length f plotted as a function of $k = 6{\cdot}85nI/\sqrt{v}$, representing the power of the lens.

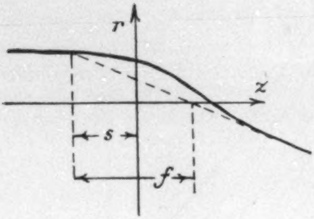


Figure 4.

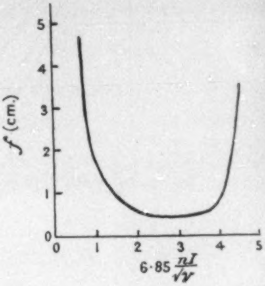


Figure 5. Variation of focal length with coil strength.

§ 5. CALCULATIONS FOR SATURATED LENS

The above results have been obtained by assuming that the permeability coefficient is greater than 1,000. This assumption is not valid for strong lenses, and its range of validity can be determined by finding the relation between the intensity of magnetization in any part of the iron and the number of ampere turns in the coil. In the lens under consideration it was found that saturation would occur if the number of ampere turns exceeded 1,200, and since the optimum strength of the lens has been found experimentally to lie between 4,000 and 5,000 ampere turns, it is evident that under normal working conditions the core and pole pieces are likely to be saturated. The possibility of using the relaxation method to calculate the field of a saturated lens was therefore examined.

If the iron in the field is saturated, the ratio B/H is not constant, and the relation $\mathbf{B} = \mu\mathbf{H}$ has to be replaced by

$$\mathbf{B} = f(\mathbf{H}),$$

where f is a function defined by the magnetization curve for the material. In an isotropic medium $\mathbf{H}$ is parallel to $\mathbf{B}$, so that

$$H_r/B_r = H_z/B_z.$$

If ψ is defined in terms of $\mathbf{B}$ as before, no simple equation can be found which is satisfied by ψ. It may, however, be assumed that the ratio B/H varies slowly over the field, and that in the region of any mesh point there is a linear relation between B and H. The coefficient μ_d is then defined as the slope of the tangent to the (B, H) curve. Thus $B - B_0 = \mu_d H$, where B_0 is the intercept of the tangent on the axis and μ_d and B_0 vary from point to point.

The function ψ may be defined as above by the equations

$$B_z = -\frac{1}{r}\frac{\partial\psi}{\partial r}, \qquad B_r = \frac{1}{r}\frac{\partial\psi}{\partial z}.$$

Then equation (2) follows as before, and from conditions (3) the boundary equations are

$$\omega_1 = \omega_2 \quad \text{and} \quad \left(\frac{\partial\omega}{\partial n}\right)_1 = \mu\left(\frac{\partial\omega}{\partial n}\right)_2,$$

where $\mu = B/H$.

The process of solution is complicated by the fact that the values of μ are not known until B has been calculated, and μ is involved in the calculation of B. A method of successive approximation is therefore necessary. Bearing in mind the physical conditions of the problem, a distribution of μ along the boundary of the iron may be assumed and the flux function calculated as before, but using the exact expressions for the boundary conditions for finite μ. Only the equations for the boundary points are affected by the variation of μ; those for the points inside the iron are exactly the same as for points in air (see Appendix). Given the flux function, B can be derived from equations (1), and hence new values of μ from the (B, μ) curve (Figure 6). After some experience of the behaviour of the distribution the computer can make the process converge fairly quickly.

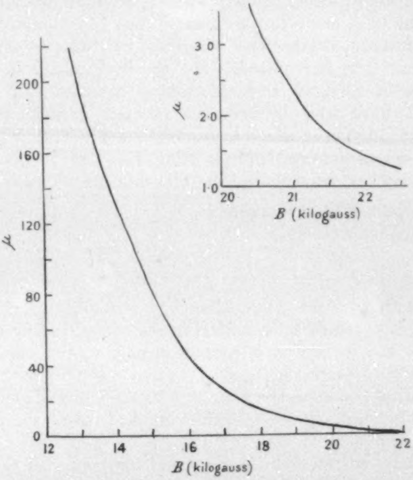


Figure 6. Variation of magnetic permeability with induction.

In order to investigate the field distribution when the pole pieces are saturated, the calculation was carried out for a lens similar to that illustrated by Hillier and Ramberg (1947, Figure 21) and used by them in attaining a resolution of 10 A. The distribution for unsaturated iron was first calculated, the field being obtained inside the iron as well as in air. Table 1 shows the distribution in the coil windings and Table 2 that in the region of the pole pieces on a larger scale. The field inside the iron in Table 2 has been obtained by assuming a uniform distribution of flux lines across the sections marked AA′ and BB′ in Table 1, these sections being taken at a sufficient distance from the pole pieces to avoid their disturbing effect on the uniformity. Table 3 shows the axial field distribution in more detail. The values of ω in these figures are given in arbitrary units; the units in Table 3 happen to be larger than those in Tables 1 and 2 by a factor 435/20. This calculation with infinite permeability showed that there is an area in the air gap where the flux lines are straight and parallel to the axis of the coil, that is, the line $\omega=20$ in Table 2. In such a region the radial boundary equation (see Appendix) with $\omega_1=\omega_3$ and $Z_0=0$ reduces to

$$\left(1-\frac{h}{2r}\right)\omega_2+\left(1+\frac{h}{2r}\right)\omega_4-4\omega_0=0,$$

which is independent of μ. Therefore the introduction of a finite value of μ in this region does not affect the distribution, and it was assumed that saturation at the outer boundaries of the pole pieces could be regarded as having no effect on the region near the axis, and that the distribution of ω along lines surrounding the saturated tips was unchanged. Thus the problem was reduced to one with definite boundary values derived from the unsaturated lens. These boundary values are shown on the dotted line in Table 4. Nothing is assumed here about the absolute value of ω along this boundary; this is undoubtedly changed by saturation at the outer edge of the iron; it is merely the constancy of the distribution along the boundary that is required. The absolute value is determined by Z, which is involved in the conversion of ω into ψ (equation (6)). Since nothing is known about Z in this case, a different method has to be adopted in deriving the intensity from the function ω, which means that there is an uncertainty about the coil strength for which the calculation is being done until it is completed.

The process can best be explained by following the calculation step by step:

1. B_z, B_r are defined by the relations

$$B_z=\frac{K}{r}\frac{\partial\omega}{\partial r},\qquad B_r=-\frac{K}{r}\frac{\partial\omega}{\partial z},\qquad \ldots\ldots(10)$$

where $K\omega=\psi$, and K in this calculation is chosen arbitrarily. The order of magnitude of K is obtained by considering the corresponding constant in the unsaturated case and increasing it proportionately to the number of ampere turns for which the calculation is required.

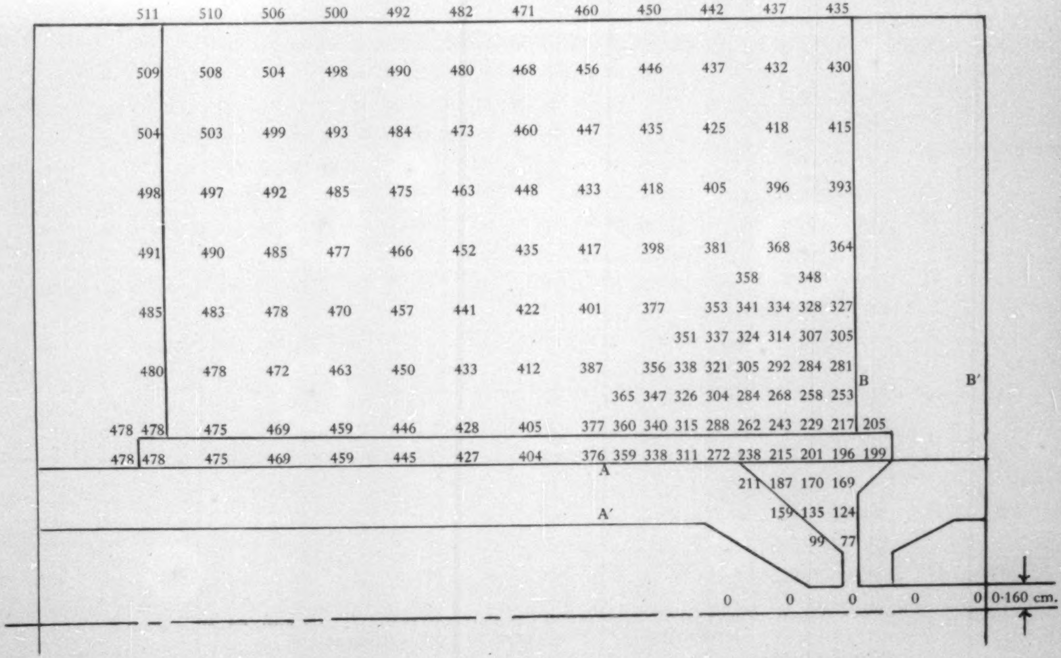
2. A distribution of μ along the boundary is then guessed, and the function ω calculated over the field by the relaxation method, using the finite difference equations in the Appendix.

3. $\partial\omega/\partial z$ and $\partial\omega/\partial r$ can then be calculated for points on the boundary by differences, and hence B_z and B_r from equations (10). Then $B^2=B_z^2+B_r^2$, and a second approximation to μ at each point can be read off the (B,μ) curve.

4. Steps (2) and (3) are repeated until B and μ correspond within the limits of accuracy required.

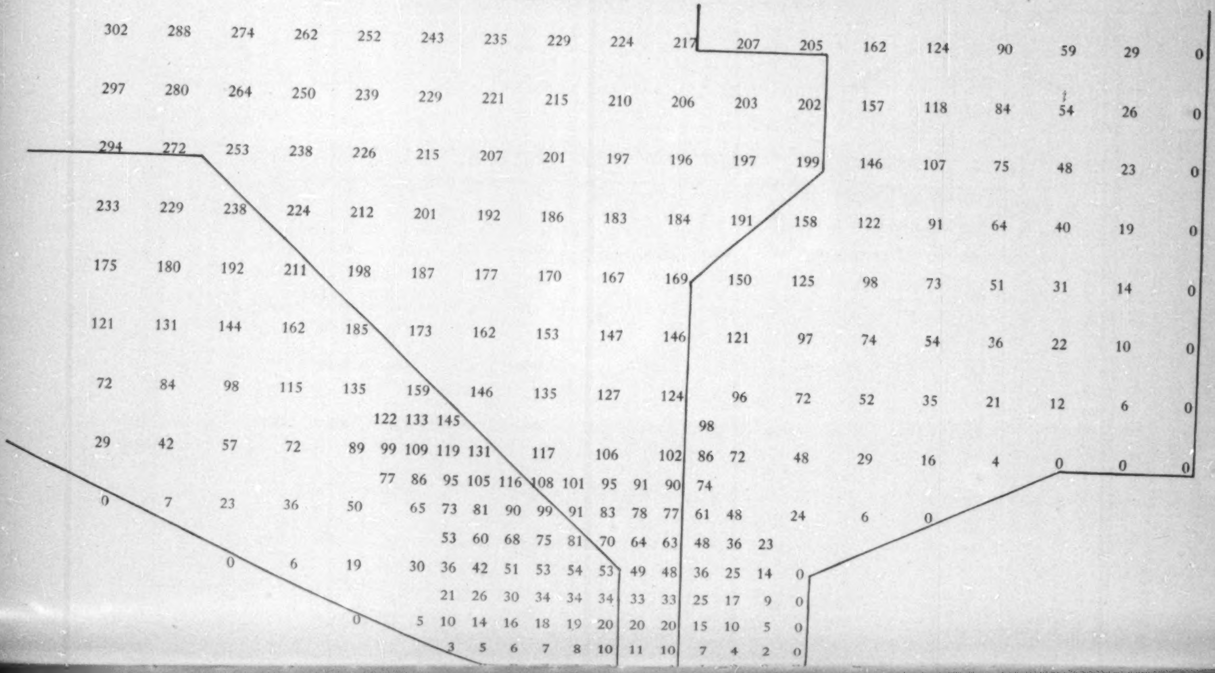
The results for one coil strength are shown in Table 4.

Table 1. Field distribution in coil windings with infinite permeability.
The numerical values are those of ω. $\psi = 0{\cdot}0162 nI . \omega$



	511	510	506	500	492	482	471	460		450		442		437		435			
	509	508	504	498	490	480	468	456		446		437		432		430			
	504	503	499	493	484	473	460	447		435		425		418		415			
	498	497	492	485	475	463	448	433		418		405		396		393			
	491	490	485	477	466	452	435	417		398		381		368		364			
													358		348				
	485	483	478	470	457	441	422	401		377		353	341	334	328	327			
											351	337	324	314	307	305			
	480	478	472	463	450	433	412	387		356	338	321	305	292	284	281			
									365	347	326	304	284	268	258	253			
478	478	475	469	459	446	428	405	377	360	340	315	288	262	243	229	217	205		
478	478	475	469	459	445	427	404	376	359	338	311	272	238	215	201	196	199		
													211	187	170	169			
														159	135	124			
															99	77			
												0		0		0		0	0

Table 2. Field distribution in the region of the pole pieces with infinite permeability. The numerical values are those of ω. $\psi = 0{\cdot}0162 nI \,.\, \omega$



302	288	274	262	252		243		235		229		224		217		207		205	162	124	90	59	29	0
297	280	264	250	239		229		221		215		210		206		203		202	157	118	84	54	26	0
294	272	253	238	226		215		207		201		197		196		197		199	146	107	75	48	23	0
233	229	238	224	212		201		192		186		183		184		191		158	122	91	64	40	19	0
175	180	192	211	198		187		177		170		167		169		150		125	98	73	51	31	14	0
121	131	144	162	185		173		162		153		147		146		121		97	74	54	36	22	10	0
72	84	98	115	135		159		146		135		127		124		96		72	52	35	21	12	6	0
					122	133	145								98									
29	42	57	72	89	99	109	119	131		117		106		102	86	72		48	29	16	4	0	0	0
					77	86	95	105	116	108	101	95	91	90	74									
0	7	23	36	50		65	73	81	90	99	91	83	78	77	61	48		24	6	0				
							53	60	68	75	81	70	64	63	48	36	23							
		0	6	19		30	36	42	51	53	54	53	49	48	36	25	14	0						
							21	26	30	34	34	34	33	33	25	17	9	0						
				0		5	10	14	16	18	19	20	20	20	15	10	5	0						
							3	5	6	7	8	10	11	10	7	4	2	0						

Table 3. Field distribution along the axis with infinite permeability, valid for

Table 3. Field distribution along the axis with infinite permeability, valid for any coil strength less than 1,000 ampere turns.

The distribution is symmetrical about OY.

The numerical values are those of ω. $\psi = 0{\cdot}000744nI\,.\,\omega$

																								Y
																			435		435		435	
																			320		320		320	
																			225		225		226	
																			182	183				
																			143	144	145		147	
																			106	108	111	113		
		0			1		1		2		4		6		12	17	24	36	64	76	82		85	
																16	22	31	43	53				
*	1		2		5		11		18		30		54		99		180	239	306		402		431	
*																		168						
†			11		28		56		99		171		300		522		867		1298		1629		1760	
†		3	7	12	17	24	33	45	60	79	104	137	181	238	308	395	496	608	720	819	896	944	960	
†		2	4	6	8	11	15	21	28	37	49	64	84	109	141	179	223	270	317	358	390	410	417	
†		0·5	1	1	2	3	4	5	7	10	13	17	22	28	36	45	56	67	78	88	96	101	102	
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
																								O

This inset shows the region of the corner enlarged

				143	143	144
				124	125	
				106	107	108
				86	89	
24	29	36	46	64	72	76
	28	34	42	52	59	
22		31		43		53

* The values in these lines are multiplied by 10. † The values in these lines are multiplied by 100.

The distribution of ω along the axis is then known, and since K has been arbitrarily chosen, the distribution of ψ is also known, and the value of nI for which the calculation has been done can be found from Ampère's law as follows:

From (8) $$h^2 H_{r=0} = h^2 K \left(\frac{\partial^2 \omega}{\partial r^2}\right)_0 = 2K\omega(h).$$

Thus $$\int H_z\, dz = \int \frac{2K}{h^2}\, \omega(h)\, dz = \frac{2K\omega_{\mathrm{m}} A}{h},$$

where $$A = \int \frac{\omega(h)}{\omega_{\mathrm{m}}} \frac{dz}{h}$$

and ω_{m} is the value of $\omega(h)$ at $z=0$.

But $$\int H_z\, dz = 0{\cdot}4\pi nI,$$

so that $$nI = \frac{K\omega_{\mathrm{m}} A}{0{\cdot}2\pi h}.$$

When one calculation has been done for the highest value of nI required, the other distributions can be obtained easily by choosing the initial approximations in the appropriate proportions. It was also sometimes useful to vary K in the middle of a calculation in order to reduce the residuals as rapidly as possible. The resulting coil strengths are shown in the table below.

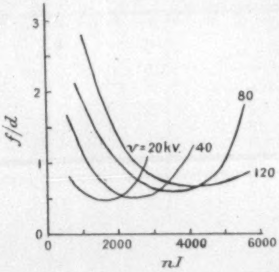


Figure 7. Variation of focal length with coil strength and voltage.

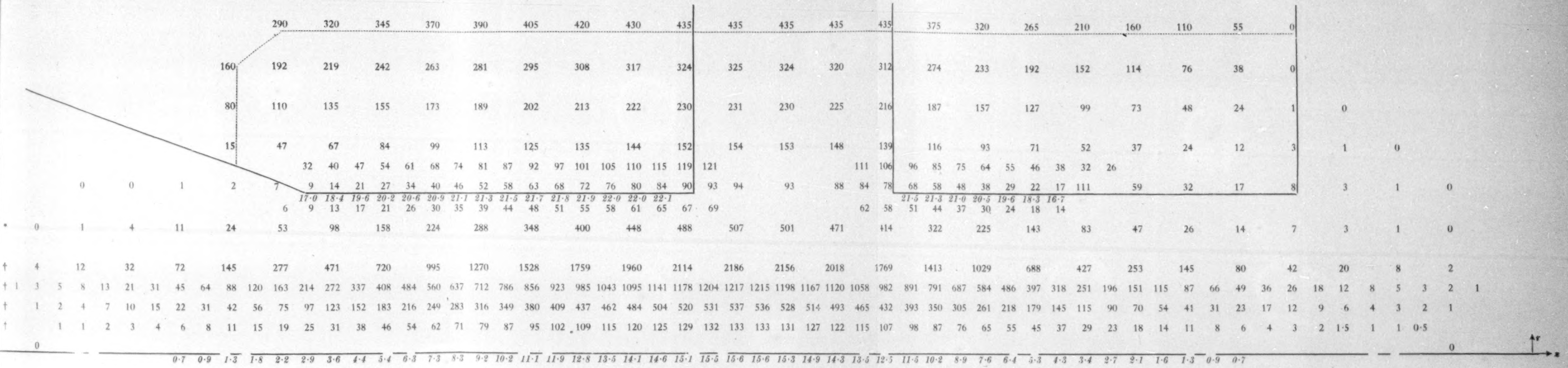
K	1·12	1·68	1·95	2·18	2·35
ω_{m}	1·045	1·14	1·20	1·29	1·31
A	12·63	16·91	18·78	20·63	22·40
H_{m} (gauss)	5850	9570	11710	14090	15350
nI (amp. turns)	1180	2580	3560	4620	5480

When the coil strength was less than 1,000 ampere turns, it was found that the effects of saturation could be neglected, and, therefore, the distribution shown in Table 3 is valid for any coil strength below this limit.

The focal lengths of the lens for various coil strengths and accelerating voltages were calculated by the method of §4, and Figure 7 shows the curves of focal length plotted against ampere turns.

Table 4. Field distribution along the axis with a coil strength of 5,480 ampere turns. The numerical values in italics are those of the total induction B, in kilogauss, the others are those of ω.

$$B_z = -\frac{K}{r}\frac{\partial\omega}{\partial r}, \quad B_r = \frac{K}{r}\frac{\partial\omega}{\partial z}, \quad \text{where } K = 2{\cdot}35.$$ The length of the small mesh is 0·020 cm.



290 320 345 370 390 405 420 430 435 435 435 435 435 375 320 265 210 160 110 55 0

160 192 219 242 263 281 295 308 317 324 325 324 320 312 274 233 192 152 114 76 38 0

80 110 135 155 173 189 202 213 222 230 231 230 225 216 187 157 127 99 73 48 24 1 0

15 47 67 84 99 113 125 135 144 152 154 153 148 139 116 93 71 52 37 24 12 3 1 0

32 40 47 54 61 68 74 81 87 92 97 101 105 110 115 119 121 111 106 96 85 75 64 55 46 38 32 26

0 0 1 2 7 9 14 21 27 34 40 46 52 58 63 68 72 76 80 84 90 93 94 93 88 84 78 68 58 48 38 29 22 17 111 59 32 17 8 3 1 0

17·0 18·4 19·6 20·2 20·6 20·9 21·1 21·3 21·5 21·7 21·8 21·9 22·0 22·0 22·1 21·5 21·3 21·0 20·5 19·6 18·3 16·7

6 9 13 17 21 26 30 35 39 44 48 51 55 58 61 65 67 69 62 58 51 44 37 30 24 18 14

* 0 1 4 11 24 53 98 158 224 288 348 400 448 488 507 501 471 414 322 225 143 83 47 26 14 7 3 1 0

† 4 12 32 72 145 277 471 720 995 1270 1528 1759 1960 2114 2186 2156 2018 1769 1413 1029 688 427 253 145 80 42 20 8 2

† 1 3 5 8 13 21 31 45 64 88 120 163 214 272 337 408 484 560 637 712 786 856 923 985 1043 1095 1141 1178 1204 1217 1215 1198 1167 1120 1058 982 891 791 687 584 486 397 318 251 196 151 115 87 66 49 36 26 18 12 8 5 3 2 1

† 1 2 4 7 10 15 22 31 42 56 75 97 123 152 183 216 249 283 316 349 380 409 437 462 484 504 520 531 537 536 528 514 493 465 432 393 350 305 261 218 179 145 115 90 70 54 41 31 23 17 12 9 6 4 3 2 1

† 1 1 2 3 4 6 8 11 15 19 25 31 38 46 54 62 71 79 87 95 102 109 115 120 125 129 132 133 133 131 127 122 115 107 98 87 76 65 55 45 37 29 23 18 14 11 8 6 4 3 2 1·5 1 1 0·5

0 0

0·7 0·9 1·3 1·8 2·2 2·9 3·6 4·4 5·4 6·3 7·3 8·3 9·2 10·2 11·1 11·9 12·8 13·5 14·1 14·6 15·1 15·5 15·6 15·6 15·3 14·9 14·3 13·5 12·5 11·5 10·2 8·9 7·6 6·4 5·3 4·3 3·4 2·7 2·1 1·6 1·3 0·9 0·7

* The values in this line are multiplied by 10.

† The values in this line are multiplied by 100.

As before, the accuracy of the relaxation calculation to the number of figures required was ensured by halving the mesh length until the distribution remained unaltered by further increase in the number of nodal points. This also justified the use of discrete values of the variable μ at boundary points. The final accuracy of the results is limited by the knowledge available about the magnetic properties of the iron, but the value of μ was not found to be very critical, especially in regions where ω did not vary rapidly. A variation in μ of the order of 10% usually only produced a variation of 1% in ω. Thus if other sources of error are neglected, the error in ω is not greater than 1%, and it is possible to find the focal length to an accuracy of 0·1 mm.

§ 6. DISCUSSION OF RESULTS

Comparison has been made above with the 'classic' field distributions of Glaser and Ramberg. Cosslett (1946) has based a discussion of the resolving power of the electron microscope on their work, and it is of interest to compare this with the results obtained by the relaxation method.

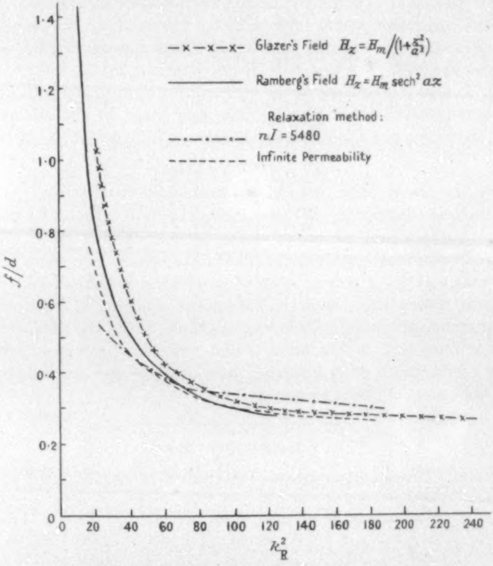


Figure 8. Variation of focal length with lens strength.

The curves of f/d against a parameter representing coil strength have been derived for both these distributions and are reproduced in Figure 8. The focal length f_R in this case has been calculated according to a different definition from that given above. The alternative definition is given by Ramberg in terms of a trajectory initially parallel to the axis, and the slope of the trajectory at the point where it cuts the axis. Thus $f_R = r_0/\Delta_0$, where $\Delta_0 h$ is the slope at the point $r=0$ (Figure 4).

As can be seen from Figure 4, the two definitions are equivalent when the focal point is in field-free space. For strong fields f_R decreases asymptotically to zero, while the geometrical focal length becomes infinite, indicating the existence of two or more focal points for fields of this value. Ramberg's focal length is more convenient for the determination of the spherical aberration coefficient, although it assumes that the part of the field beyond the focal point does not contribute appreciably to the aberration. This assumption is justified in practice, since lenses are not worked at strengths beyond the minimum geometrical focal length, and in the working region the two focal lengths do not differ greatly.

In Figure 8 the abscissa k_R^2 is the lens parameter as defined by Ramberg, and is given by

$$k_R^2 = d^2 H_m^2/(2{\cdot}63)^2 V.$$

The Figure also shows the corresponding curves obtained by the relaxation method for the case of infinite permeability and for a coil strength of 5,480 ampere turns. The agreement is close, especially for values of k_R^2 between 60 and 100, and indicates that although the field distributions represented by Glaser's and Ramberg's expressions are not good approximations to the field of a practical lens, they do give a good idea of the variation of lens power with coil strength and accelerating voltage, and give a satisfactory basis for the calculation of the spherical aberration and the discussion of lens performance.

§7. CONCLUSIONS

During the fifteen years or so of the development of electron microscopes, progress in their construction has been made chiefly by trial-and-error methods, guided by mathematical treatment where approximations made this possible. A great many empirical generalizations have been made as a result of this work, and optimum operating conditions are now almost standardized. In view of the constant testing that these generalizations undergo in practice, it seems unlikely that the application of relaxation methods to the standard designs would lead to anything new at this stage. The relaxation method might however provide a useful means of investigating new designs, and one which avoids the practical difficulties of direct measurement.

APPENDIX

The finite difference equations to be satisfied at various types of points are as follows:

Ordinary points

$$\omega_1 + \left(1 - \frac{h}{2r}\right)\omega_2 + \omega_3 + \left(1 + \frac{h}{2r}\right)\omega_4 - 4\omega_0 + h^2 Z_0 r = 0.$$

Irregular stars

$$\frac{2\omega_1}{x_1(x_1+x_3)} + \frac{1}{x_2(x_2+x_4)}\left(2 - \frac{hx_4}{r}\right)\omega_2 + \frac{2\omega_3}{x_3(x_1+x_3)} + \frac{1}{x_4(x_2+x_4)}\left(2 + \frac{hx_2}{r}\right)\omega_4 - \left\{\frac{2}{x_1x_3} + \frac{2}{x_2x_4} + \frac{h}{r}\frac{x_2 - x_4}{x_2x_4}\right\}\omega_0 + h^2 Z_0 r = 0.$$

Diagonal net—ordinary points

$$\left(1-\frac{h}{2r}\right)(\omega_1+\omega_2)+\left(1+\frac{h}{2r}\right)(\omega_3+\omega_4)-4\omega_0+2h^2Z_0r=0.$$

Diagonal net—irregular stars

$$\frac{1}{x_1(x_1+x_3)}\left(1-\frac{hx_3}{2r}\right)\omega_1+\frac{1}{x_2(x_2+x_4)}\left(1-\frac{hx_4}{2r}\right)\omega_2+\frac{1}{x_3(x_1+x_3)}\left(1+\frac{hx_1}{2r}\right)\omega_3$$
$$+\frac{1}{x_4(x_2+x_4)}\left(1+\frac{hx_2}{2r}\right)\omega_4$$
$$-\left\{\frac{1}{x_1x_3}\left(1+\frac{hx_1}{2r}-\frac{hx_3}{2r}\right)+\frac{1}{x_2x_4}\left(1+\frac{hx_2}{2r}-\frac{hx_4}{2r}\right)\right\}\omega_0+h^2Z_0r=0.$$

Radial boundary points (Figure 9)

$$2\omega_1+(\mu+1)\left(1-\frac{h}{2r}\right)\omega_2+2\mu\omega_3+(\mu+1)\left(1+\frac{h}{2r}\right)\omega_4-4(\mu+1)\omega_0+h^2Z_0r=0.$$

Cylindrical boundary points (Figure 10)

$$\left(1-\frac{h}{2r}+\frac{h\mu}{2r}\right)\omega_1+2\left(1-\frac{h}{2r}\right)\omega_2+\left(1-\frac{h}{2r}+\frac{h\mu}{2r}\right)\omega_3+2\mu\left(1+\frac{h}{2r}\right)\omega_4-$$
$$4\left(1-\frac{h}{2r}+\frac{h\mu}{2r}\right)\omega_0+\left(1+\frac{h}{2r}\right)h^2Z_0r=0.$$

Right-angled corner (Figure 11)

$$\omega_1+\left(2\mu-1-\frac{h}{2r}\right)\omega_2+(2\mu-1)\omega_3+\left(1+\frac{h}{2r}\right)\omega_4-4\omega_0+h^2\mu Z_0r=0.$$

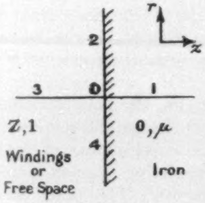


Figure 9.

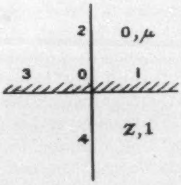


Figure 10.

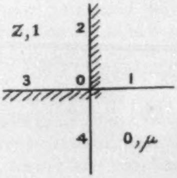


Figure 11.

ACKNOWLEDGMENTS

The author wishes to thank Professor G. I. Finch for permission to work in his laboratory and for his interest and advice during the course of the work, and also Mr. D. N. de G. Allen for assistance on several occasions.

REFERENCES

Bertram, S., 1940, *Proc. Inst. Radio Engrs., N.Y.*, **28,** 418; 1942, *J. Appl. Phys.*, **13,** 496.
Cosslett, V. E., 1946, *Proc. Phys. Soc.*, **58,** 443.
Glaser, W., 1941, *Z. Phys.*, **117,** 285.
Hillier, J., and Ramberg, E. G., 1947, *J. Appl. Phys.*, **18,** 48.
Ramberg, E. G., 1942, *J. Appl. Phys.*, **13,** 582.
Southwell, R. V., 1946, *Relaxation Methods in Theoretical Physics* (Oxford: University Press).

An Electrostatic Fluxmeter of Short Response-time for use in Studies of Transient Field-changes

By D. J. MALAN and B. F. J. SCHONLAND

The Bernard Price Institute, University of the Witwatersrand, Johannesburg

MS. received 24*th October* 1949, *and in final form* 26*th January* 1950

ABSTRACT. The fluxmeter consists of a conducting system of small capacity which is alternately exposed to and screened from the electric field by the movement of a rapidly rotating metal shield, thus producing an alternating E.M.F. of period 0·83 millisecond. The output, after amplification, is displayed on the screen of a cathode-ray oscillograph and the record carries an automatic indication of the sense of the field every 7·5 milliseconds.

At the maximum sensitivity employed the device gives a deflection of one centimetre in a field of 20 volts/metre, with a background noise-level of 0·15 cm. (3 volts/metre).

The response-time of the instrument is one half-cycle (0·42 millisecond).

The fluxmeter has been developed to give the rapid response and high sensitivity required for studies of the electric fields of thunderstorms in the intervals between the separate strokes of a lightning discharge.

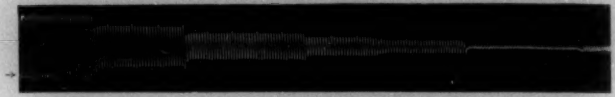
§ 1. INTRODUCTION

THE electrostatic fluxmeter described in this paper has been developed to provide a convenient method of measuring the electric fields produced by thunderstorms. The principle of the device was first described by Ross Gunn (1932) and it is sometimes referred to as a 'field-mill'. It does not appear to have been developed before in a form which gives it either the sensitivity or the rapid time of response of the instrument to be described here. It has been used by us in studies of the electric fields of thunderstorms in the interval of the order of 0·03 second between separate component strokes, and it has an obvious application in providing continuous records of electric fields over long periods. With a time resolution of somewhat less than one millisecond, it is capable of resolving field changes which the beautiful capillary electrometer method of C. T. R. Wilson (1916) misses, and in our experience it is more suitable for field strength records of long duration than the combination of exposed conductor and cathode-ray oscillograph devised by Appleton, Watson-Watt and Herd (1926), particularly when studying nearby storms. The latter method calls for high insulation resistance of the exposed conductor in order to maintain a long time-constant and this is a matter of some difficulty in the presence of heavy rain and point-discharge.

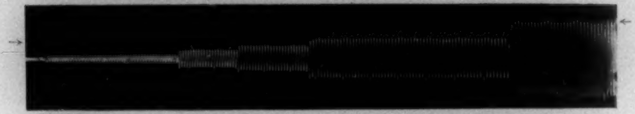
§ 2. DESCRIPTION OF THE INSTRUMENT

The instrument consists essentially of a conducting system of small capacity which is alternately and regularly exposed to, and screened from, the electric field by the movement of a rapidly rotating earthed metal shield. The exposure-screening cycle is carried out 1,200 times per second.

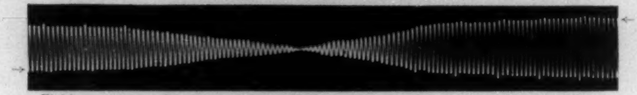
The conditions under which the cycle is performed cause an amplifying system connected to the underside of the conductor to give rise to an alternating E.M.F. of approximately sinusoidal form, whose amplitude is proportional to the strength of the field to be measured. This E.M.F. is amplified and passed to the X plates of a cathode-ray tube and the waveform is recorded by photographing the screen



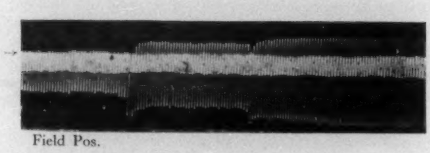
Field Neg. SBM 1·3



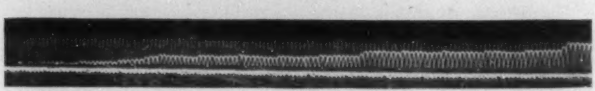
Field Pos. continuation of above



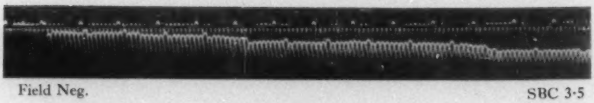
Field Neg. Zero Pos. SBM 1·4



Field Pos. SAX 2·6



Field Pos. SBC 4·1



Field Neg. SBC 3·5

Figure 1.
Time scale: 1 cycle=0·8 millisec.

To face page 408

in a drum-film camera rotating about a horizontal axis. The amplitude of the envelope of the crests and troughs on this record is proportional to the field strength and alters as the field strength alters. A special device described in § 5 indicates the sense of the field every ninth cycle and so shows which side of the envelope is to be selected for measurement. Some examples of records taken with the equipment using two different methods of recording are given in Figure 1 (see Plate). In all of these if the crests are enhanced in amplitude by 'polarity pips' the field is negative and vice versa. The arrows, therefore, show the side of the envelope which is used for measurement.

Constructional details of the instrument are shown diagrammatically in Figure 2, which is, however, not drawn to scale.

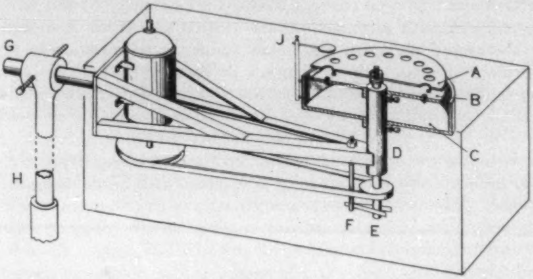


Figure 2.

The conducting system is formed by 18 metal studs mounted on a paxolin disc B. These studs, four of which are shown in the diagram, are constructed from No. 4 B.A. cheese-head screws of diameter 6·5 mm. and are set in a circle of diameter 20 cm. They are connected together below in the cylindrical brass box C and joined to the control grid of a low capacity triode valve V_1 (not shown) which is mounted inside C. The output of V_1 is led by cable to such further amplifiers as are necessary and from them to the cathode-ray tube. The capacity of the stud-grid system is 50 pF.

The screening disc A is of brass 23·5 cm. in diameter and 1·3 mm. thick. It is provided with 18 equally spaced holes of diameter 2·0 cm. whose centres lie immediately above the studs. The spacing between A and B can be adjusted by altering the position of B after loosening its clamping screws and is usually 2 mm. A is rotated at 4,000 r.p.m. by the steel shaft E, mounted on ball-bearings in the tubular housing D, and carrying a pulley which is belt-driven by a $\frac{1}{4}$ H.P., 1,400 r.p.m. A.C. motor.

The box C is mounted, like B, on the bearing housing D by means of a collar and clamping screws and by moving it downwards one can get at the valve without dismantling the two discs. In order to provide the best possible shielding for this valve, the box is provided with an earthed cover in the form of a thin metal sheet attached to the top of disc B and suitably perforated to give insulation to the studs.

The framework supporting the device is built of angle-iron with welded joints and carried on a horizontal pipe G, which enters a socket at the top of a vertical pipe H, one metre long, mounted on the roof parapet of the laboratory.

The whole framework is enclosed in a rectangular earthed metal case with the A side open and usually upwards. When it is necessary to run the instrument in rain or hail, it is inverted, so that A is downwards, and it is swung out over the outside wall of the laboratory so as to obtain a suitable ground clearance. Its sensitivity in the inverted position is reduced by a factor of two.

An essential part of the equipment is the earthing brush F, which consists of two springy phosphor-bronze strips pinching the shaft E between them. The surface of contact is slightly oiled and requires periodic cleaning. The purpose of F is to get rid of frictional charges which collect in an irregular manner on the disc A as a result of its rotation and of the belt drive. The brush and the whole framework are connected to earth. Unless this earthing is very good the noise-level of these frictional charges swamps the output.

To obtain a smooth output from the device, the discs A and B must be carefully constructed and quite flat. Any tendency to vibration in A must be eliminated since it gives rise to changes in capacity and so to an unwanted fluctuation in the output.

§ 3. AMPLIFYING CIRCUITS

Figure 3 shows the arrangement used in the head-amplifying circuit of V_1. In order to obtain a low grid capacity, a Mullard E.F.37 pentode is connected up as a triode. This valve was chosen for its low microphonic properties, and as an extra precaution it is suspended by rubber bands. The effect of the grid resistor R_1 upon sensitivity is discussed in § 7. The output from V_1 is taken from a cathode-follower circuit and is passed to the recording room through a screened cable 7 metres long. The H.T. and filament supplies for V_1 are also provided through screened cable from the recording room and the whole equipment can thus be operated from a distance.

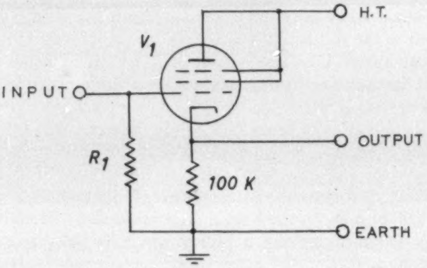


Figure 3.

Further amplification as required is provided by the X plate amplifier of the Dumont 208B oscillograph used to display the final output.

§ 4. METHODS OF RECORDING

In the present work we have been concerned with the recording of field changes associated with lightning flashes and have been able to use special trigger methods and fairly high film speeds. The drum camera used by us has a peripheral velocity of 44 cm/sec.

Two systems of recording have been employed. In the first, the time-base of the oscillograph is not used and the output is seen on the tube face as a horizontal line, which is screened off by a thin strip of black paper. An anticipatory lightning trigger (Schonland and Elder 1941) is arranged to operate in the initial stages of the stepped leader process of a flash to ground and to put a bias on the Y plates of the oscillograph before the leader has travelled very far. This bias causes the spot to move out from behind the screening strip and is kept on for one second so that the field changes taking place during the flash can be recorded. The method has the disadvantage that information about early first-leader field changes is lost, though it can be provided, if required, by another oscillograph kept continuously running. The first four records of Figure 1 are examples of this type of recording.

In the second method no trigger is used but the adjustable time-base of the oscillograph is employed. This time-base is 'locked' to the alternating E.M.F. from the studs so that the unmasked screen would show a single stationary wave pattern whose axis is vertical and whose amplitude alters with the field. A paper mask with a narrow horizontal slit is attached to the face of the screen in such a way as to cut out all but the crest of this stationary wave. When the field is steady or is varying slowly, the crest can be kept hidden behind a small masking disc at the centre of the slit, by manual adjustment of the X shift of the oscillograph. When the field alters rapidly, the crest travels along the slit and its displacement is recorded on the drum camera. In the case of many lightning flashes the sudden changes of field due to the leader process initiate the movement required for recording.

In Figure 1, SBC 4·1 and SBC 3·5 are examples of this type of recording; in the first the crest-spot has moved to a position above the masking disc, in the second it has moved below it.

§ 5. POLARITY INDICATOR

In order to show directly the sense or polarity of the electric field whose strength, as explained, is measured by the envelope of the amplitudes of the waveform, we have incorporated in the instrument a device which automatically increases the height of the crests of the waveform every ninth cycle if the field is negative in sense and the troughs every ninth cycle if it is positive. An inspection of the record immediately gives the required sense of the field.

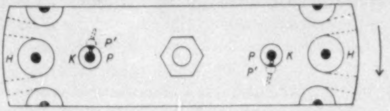


Figure 4.

Figure 4, which is a plan view of a portion of the disc A of Figure 2, shows how this is done. HH are the exposure holes in the screening disc, with the studs of disc B momentarily below their centres. An extra pair of holes, KK, have been placed in the positions shown. Below each of these are two extra studs P and P′ fixed in slots in the paxolin disc B. Thin wires join these four 'polarity studs' to the output lead of the main studs. The phase of their contribution to the output can be adjusted by moving them in their slots.

When the disc A rotates in the direction of the arrow, the leading edges of the holes K and H uncover the peripheral studs and the studs P at the same moment, thus enhancing the output by the contributions from P. In the absence of studs P′, both crests and troughs will be larger than before. The P′ studs, however, become exposed when all other studs are being screened, so that their output is opposed in sense to that of the other studs. By adjustment of the position of the P′ studs their output can be arranged to neutralize that of the P studs during the half-cycle that they are exposed. In this way the 'exposure' amplitude peaks of the output are larger than normal every half-turn of disc A but the 'screening' troughs are not affected. On reversal of the sense of the external field, the exposure half-cycle is that of the troughs and the polarity indication, as shown in the example marked SBM 1·4 (Figure 1) promptly shifts over to the other side of the zero line. All the records in Figure 1 show this polarity indication clearly. The easiest to follow is SBM 1·4 which begins with a negative electric field as shown by the enhanced crests and then passes through zero to an almost equal positive value, as shown by the enhanced troughs. This reversal of field occupies 0·06 second.

SBM 1·3, as shown by its enhanced crests at the start, begins with a large negative field which falls to nearly zero after five abrupt positive field changes caused by successive strokes of a lightning flash. In the record below the first the field is positive and five more strokes lead to a large positive field which ends the discharge. The sense of the fields in the other three examples shown in Figure 1 is similarly identified.

In practice it is necessary to adjust the height of the P studs with thin packing washers to bring them nearer to A and so to give them a slightly larger output than the P′ studs.

§6. SENSITIVITY AND PERFORMANCE

The instrument in the form at present used by us gives a peak amplitude of 1 cm. in a field of 20 volt/metre and thus easily registers the electric field of fine weather which is of the order of 100 volt/metre. We have been able to record with it field changes due to separate lightning strokes at a distance of 25 km. The noise-level on our present site is about 1·5 mm. at the maximum amplification used and at the worst times of day. In a quieter locality the sensitivity could be much increased by using a larger disc with larger holes and larger studs or by cutting sectors from the disc, as shown by the dotted lines in Figure 4, and mounting the studs on a circle near to the edge of the disc. Local sources of electrical interference have, however, prevented us from using a higher sensitivity.

During rain when the arrangement is inverted and swung outwards, it gives half the above sensitivity when one metre from the building in a horizontal direction and 12 metres above the ground.

The equipment is calibrated by taking readings in fine weather and comparing them with simultaneous measurements by the stretched wire method in a nearby open field.

The electric fields of nearby thunderstorms are so large as to require reduction of the amplification provided in the oscillograph if the instrument is not to go 'off-scale'. This is easily met by adjustment of the calibrated attenuator in the oscillograph. It often happens, however, that there are charged clouds in the neighbourhood when one wants to register small field changes due to distant storms and to use a high amplification. We have employed two methods to balance out such disturbing fields. The first makes use of a small insulated plate J (Figure 2), fixed to the box C and mounted above the rotating disc A. A controlled balancing potential can be applied to J to prevent the local field from driving the spot off-scale. The second method is to insulate the whole framework at the socket G of Figure 2 so that it and the disc A can be charged to an appropriate balancing potential while the studs and cathode follower are separately connected to earth as before.

Trouble was initially experienced with 50 c/s. hum from the A.C. mains but this was eliminated by balancing it against the output from a variable three-phase mains transformer.

§7. THE RESPONSE-TIME OF THE INSTRUMENT

Any sudden change ΔE in the electric field strength must produce a change $-\Delta Q$ in the charge on the upper surface of the system, given by

$$-\Delta Q = kA\,\Delta E/4\pi,$$

where k is a site-correcting factor and A is the effective surface area of the conductor. A corresponding opposite charge ΔQ must appear on the under-surface. Provided the field change occurs during an exposure interval, the initial response of a device of this kind is therefore instantaneous. It can be shown, however, that the final steady value of the amplitude is less than the initial response and that a certain interval of time is needed for it to reach this steady amplitude after a sudden change.

Let us assume for simplicity that both exposure and screening intervals are equal to T and that both operations start and stop instantaneously. If C is the capacity of the system, including the grid of V_1 and the connecting leads, and R_1 is the grid resistor of Figure 3, we define α as e^{-T/CR_1}. It is then easily shown that successive half-cycles of exposure and screening give rise to changes in the charge on the under-surface of the conductor which are given by the terms of the converging series

$$\Delta Q,\ -\Delta Q(1-\alpha),\ \Delta Q(1-\alpha+\alpha^2),\ \ldots\ \Delta Q[1-\alpha+\alpha^2\ldots+(-1)^n\alpha^n].$$

These terms converge to a final steady value $\Delta Q/(1+\alpha)$ when n is large, so that the final steady amplitude is $1/(1+\alpha)$ of that obtained from the first exposure 'kick', ΔQ.

In our usual arrangement, T is 0·40 millisecond, C is 50 pF. and R_1 is 2 megohms. α is then 0·0183 and the final steady amplitude is 1·8% less than the first 'kick'.

The response-time of the instrument will be defined as the number of half-cycles, N, before the amplitude reaches 99% of its final steady value. Since the nth term in the above series can be written

$$[1+(-1)^n\alpha^{n+1}]\,\Delta Q/(1+\alpha)$$

the number N is given by $\alpha^{N+1} \simeq 0{\cdot}01$.

For the usual arrangement employed by us we have $\alpha^1 = 0{\cdot}0183$, $\alpha^2 = 0{\cdot}0033$ so that N should be one half-cycle.

Though the actual conditions of operation depart from the simple assumptions made in deriving this criterion, the instrument does, in fact, give a response-time of one half-cycle with $R_1 = 2$ megohms. With a period of alternation of 0·83 millisecond,and a response-time of less than one complete cycle, the device therefore satisfies the requirements for which it was designed.

To obtain a higher sensitivity in observing distant thunderstorms we have occasionally increased R_1 to 10 megohms. Since in this case e^{-9T/CR_1} is 0·01, the response-time should be 8 half-cycles or 3·6 milliseconds. It is found in practice to be about 9 half-cycles.

With this value of R_1, α becomes 0·449 and the final steady amplitude is only 0·69 of that given by the first half-cycle.

The amplitudes of successive half-cycles, as can be seen from the series given above, are not at first equal on the two sides of the zero. They only approach equality after the lapse of the N half-cycles of the response-time. Thus records taken with $R_1 = 10$ megohms, such as SAX2·6 of Figure 1, show an apparent shift of the zero when a sudden field change takes place. In practice this is sometimes useful in picking out small changes of field. The other records hardly show the effect, since they were taken with $R_1 = 2$ megohms and the zero shift after a field change lasts for 2 half-cycles at most.

ACKNOWLEDGMENTS

We wish to thank Dr. P. G. Gane for helpful advice on the development of this instrument, and Mr. V. Barnard and Mr. J. A. Keiller for their assistance in its construction.

REFERENCES

Appleton, E. V., Watson-Watt, R. A., and Herd, J. F., 1926, *Proc. Roy. Soc.* A, **111,** 654.
Gunn, R., 1932, *Phys. Rev.*, **40,** 307.
Schonland, B. F. J., and Elder, J. S., 1941, *J. Franklin Inst.*, **231,** 39.
Wilson, C. T. R., 1916, *Proc. Roy. Soc.* A, **92,** 555.

Wave Propagation in a Slipping Stream of Electrons : Small Amplitude Theory

By G. G. MACFARLANE and H. G. HAY

Telecommunications Research Establishment, Ministry of Supply, Great Malvern, Worcs.

MS. received 8th December 1949

ABSTRACT. A theoretical study is presented of the TM-waves that can travel along a slipping stream of electrons. A slipping stream is defined as one in which the electrons move in parallel paths with velocity which varies with distance transverse to the motion. It is found that amplifying waves can travel along a slipping stream for all frequencies. It is also found that the slipping-stream tube has the remarkable property of combining the characteristics of a two-beam tube and a travelling-wave tube. This is due to the occurrence in the stream under suitable conditions of resonance layers, which act as highly reactive impedance sheets and can guide TM-waves of slow phase velocity in the same way as the spiral or corrugated surface in a travelling-wave tube.

The slipping-stream tube behaves differently according to the fractional velocity range of the electrons. If the velocity varies linearly from V_{-a} to V_a across the stream and $\alpha=(V_a-V_{-a})/(V_a+V_{-a})\lesssim 0{\cdot}42$ it behaves primarily as a two-beam tube with plasma resonance frequency $\omega_0/2$, and therefore having a cut-off frequency $\omega_1=\omega_0/\alpha\sqrt{2}$ and maximum gain $2{\cdot}1\omega_0/V_0$ decibels per unit length, where ω_0 is the plasma resonance frequency of the slipping stream; in addition it has a low gain for $\omega>\omega_1$. However, if $\alpha>0{\cdot}42$ it behaves primarily like a travelling-wave tube and the rate of gain is about $0{\cdot}53\omega_0/V_{-a}$ decibels per unit length at all frequencies above the plasma resonance frequency. This rate of gain is low compared with the gain possible with a travelling-wave tube in which the electron velocity is V_{-a} but it is achieved without the need for any external slow-wave waveguide.

Wave propagation along a slipping stream inside a waveguide, which in the absence of the electrons can guide a TM-wave of slow phase velocity, is also discussed. It is shown that this slipping stream travelling-wave tube has very much the same characteristics as a travelling-wave tube with a uniform electron beam. Maximum gain occurs when the phase velocity v_0 of the wave in the empty guide is about equal to the velocity of the electrons nearest to the reactive impedance sheet of the guide—in the example considered this is the fastest electron velocity V_a. A parametric set of curves is given relating the complex propagation constant to the frequency for different ratios v_0/V_a.

---

§ 1. INTRODUCTION

In his original article on the electron-wave tube Haeff (1949) describes a tube in which a microwave signal can be amplified by interaction with a slipping stream of electrons. The slipping stream is a cylindrical beam in which the electrons are constrained to move in an axial direction, by the application of a strong steady axial magnetic field, with velocity which increases with radial distance from the axis of the beam. The stratification of velocity is obtained by surrounding the beam with a cylindrical conducting sheath or drift tube. The electron space charge within the drift tube depresses the potential along the axis, so that outer electrons travel faster than electrons along the axis (Smith and Hartman 1940). This results in what we shall call a slipping stream of electrons.

An analysis of the operation of this tube is not given by Haeff, although he does discuss a simplified theory of the two-beam tube in some detail. Experimental results are cited as proof that the slipping stream tube can amplify microwave signals. One of the most interesting characteristics of such a tube is the spectrum of frequencies for which amplification is possible, and this is not discussed by Haeff.

A steady slipping stream of electrons can also be maintained in an electron beam by the application of crossed electric and magnetic fields transverse to the axis of the beam. No strong axial focusing magnetic field is then required. Moreover when the electrons are perturbed by an oscillatory field they are no longer constrained to move in an axial direction but can oscillate transversely. However, to obtain parallel flow of the electrons in the non-oscillatory state the applied magnetic field must have a prescribed variation across the beam which depends on the electron density distribution, and this in turn is related to the velocity-distribution. If these values are not maintained the beam may alter in width along its length and the motion will cease to be single-stream. The non-oscillatory state of this system is the same as the single stream or type-S state of a linear magnetron in the particular example in which the electron density and the magnetic field are constant. In the more general case it would correspond to a type-S state of a magnetron with non-uniform electron density and magnetic field. In the oscillatory state, however, the periodicity of the field in the magnetron oscillator imposes the condition that the propagation constant must be real and the frequency, which is determined by the boundary conditions, must be complex for oscillations to build up in time, whereas in the slipping stream an oscillation of known real frequency impressed on the stream induces a wave to travel with increasing amplitude along the stream. The propagation constant, which is now determined by the boundary conditions, must therefore be complex. It is the slipping-stream tube with crossed electric and magnetic fields that we shall analyse in the following report.

Bunemann (1944) was the first to give a theory of the build-up of small amplitude oscillations from the type-S state in a magnetron oscillator and to derive an instability criterion. The analysis that we shall give for the slipping-stream amplifier will follow along much the same lines as that of Buneman; in particular we shall also use an action function to describe the electron flow.

One of the main objects of our work is to determine the spectrum of the frequencies which can be amplified. We shall consider not only a non-enclosed, or free, beam but also one which is loaded along its length by impedance sheets, which in the absence of the beam can guide a wave of slow phase velocity. The latter represents the beam inside a slow-wave waveguide in the manner of a travelling-wave tube. Although a loaded guide is not necessary for supporting an amplifying wave in a slipping stream of electrons it is of interest from the point of view of coupling the beam to the impressed field at the input and of extracting the amplified wave from the beam at the output. Also if ohmic losses in the walls of the guide are small it becomes apparent that a much higher rate of amplification is possible with a slipping-stream travelling-wave tube than with a free slipping stream.

§ 2. SMALL AMPLITUDE THEORY

(i) *An Arbitrary Velocity Distribution*

Our object is to investigate the properties of wave motions of small amplitude which can travel along a parallel beam of electrons, wherein the mean velocity varies with position across the stream. We use a quasi-stationary approximation throughout, in which retardation effects are neglected. For simplicity we consider only a two-dimensional variation in a cartesian coordinate system chosen so that there is no variation of any quantity in the direction of the x-axis. In the non-oscillatory state the electrons are assumed to move parallel to the z-axis with a velocity which depends on y. The stream is therefore stratified parallel to the z-axis. We shall further assume that in the non-oscillatory state the electrons are constrained to move in parallel straight paths by the combined action of a magnetic field of vector potential $(0, A_y, 0)$ and of an electric field E_y. Then if $\mathbf{V}$ is the electron velocity, ϕ the electric potential, ρ the charge density, we can obtain a non-oscillatory state for which $\operatorname{curl} m\mathbf{V} = e\mathbf{B}$ by the use of the Hamilton–Jacobi equation (see Appendix I) and Poisson's equation. For this state,

$$\left.\begin{aligned} \mathbf{V} &= [0,\ 0,\ V(y)], \\ \mathbf{A} &= [0,\ A_y,\ 0], \\ \phi &= \frac{m}{2e} V^2(y) = \phi_0, \\ \rho &= \frac{\epsilon m}{2e}\left\{V\frac{d^2V}{dy^2} + \left(\frac{dV}{dy}\right)^2\right\}. \end{aligned}\right\} \qquad \ldots\ldots(1)$$

Then an action function ψ_0 of the electrons exists, and is defined by

$$\operatorname{grad}\psi_0 = m\mathbf{V} - e\mathbf{A}.$$

Moreover, the condition $\operatorname{curl} m\mathbf{V} = e\mathbf{B}$ enables us to use the Hamilton–Jacobi equation, and is fulfilled if, as we assume in our case, the electrons leave the cathode with zero velocity and there is no component of magnetic field normal to the cathode surface (Appendix I).

Therefore

$$\left.\begin{aligned} \frac{\partial\psi_0}{\partial z} &= mV(y), \\ \frac{\partial\psi_0}{\partial y} &= -eA_y, \end{aligned}\right\} \qquad \ldots\ldots(2)$$

and the magnetic field is $[B_x, 0, 0]$, where

$$B_x = \frac{m}{e}\frac{dV}{dy}. \qquad \ldots\ldots(3)$$

The second equation of (2) defines the magnetic vector potential, from which the transverse magnetic field B_x is obtained.

It should be noted that both the charge-density distribution and the magnetic field, and hence the electric field E_y, are determined by the velocity $V(y)$.

We consider a perturbation in which $\operatorname{curl}(m\mathbf{V} - e\mathbf{A})$ is still zero. Then we specify a possible case if we satisfy

$$\left.\begin{aligned} &\operatorname{grad}\psi = m\mathbf{V} - e\mathbf{A} \\ \text{and}\qquad &e\phi = \frac{1}{2m}(e\mathbf{A} + \operatorname{grad}\psi)^2 + \frac{\partial\psi}{\partial t} \end{aligned}\right\} \qquad \ldots\ldots(4)$$

(which ensure that the equation of motion is satisfied), together with Poisson's equation and the equation of continuity. We consider a ψ of the form

$$\psi = \psi_0 + \psi_1(y)e^{i\theta},$$

where

$$\theta = \omega t - hz. \qquad \ldots\ldots(5)$$

In keeping with the restriction of the analysis to a quasi-stationary approximation, any perturbation of the magnetic vector potential will be neglected. Inserting (5) into (4) and retaining only first order perturbation terms we get

$$\phi = \phi_0 + \phi_1 e^{i\theta},$$

where

$$\phi_1 = \frac{i\omega}{e}\left(1 - \frac{hV}{\omega}\right)\psi_1. \qquad \ldots\ldots(6)$$

Poisson's equation is

$$\frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = \frac{\rho}{\epsilon}. \qquad \ldots\ldots(7)$$

We assume a small perturbation $re^{i\theta}$ on the charge density ρ. Then, from (7),

$$r = \frac{i\omega\epsilon}{e}\left[-h^2\left(1 - \frac{hV}{\omega}\right)\psi_1 + \frac{d^2}{dy^2}\left\{\left(1 - \frac{hV}{\omega}\right)\psi_1\right\}\right]. \qquad \ldots\ldots(8)$$

The next step is to eliminate r from (8) using the equation of continuity. Thus

$$\begin{aligned} J_z = \rho V_z &= (\rho + re^{i\theta})\left(V - \frac{ih}{m}\psi_1 e^{i\theta}\right) \\ &\simeq \rho V + \left(Vr - \frac{ih}{m}\rho\psi_1\right)e^{i\theta} \qquad \ldots\ldots(9) \end{aligned}$$

and

$$\begin{aligned} J_y = \rho V_y &= (\rho + re^{i\theta})\left(\frac{1}{m}\frac{\partial\psi_1}{\partial y}e^{i\theta}\right) \\ &\simeq \frac{\rho}{m}\frac{\partial\psi_1}{\partial y}e^{i\theta}, \qquad \ldots\ldots(10) \end{aligned}$$

and the equation of continuity is

$$\frac{\partial J_z}{\partial z} + \frac{\partial J_y}{\partial y} = -\frac{\partial}{\partial t}(\rho + re^{i\theta}). \qquad \ldots\ldots(11)$$

Therfore from (9), (10) and (11)

$$r\left(1 - \frac{hV}{\omega}\right) = \frac{i}{m\omega}\left\{-h^2\rho\psi_1 + \frac{d}{dy}\left(\rho\frac{d\psi_1}{dy}\right)\right\}. \qquad \ldots\ldots(12)$$

On eliminating r from (8) and (12) we get the following differential equation for ψ_1:

$$\begin{aligned} \frac{d}{dy}\left[\left\{\left(\frac{\omega_0}{\omega}\right)^2 - \left(1 - \frac{hV}{\omega}\right)^2\right\}\frac{d\psi_1}{dy}\right] + \Bigg[\frac{h}{\omega}\left(1 - \frac{hV}{\omega}\right)\frac{d^2V}{dy^2} \\ - h^2\left\{\left(\frac{\omega_0}{\omega}\right)^2 - \left(1 - \frac{hV}{\omega}\right)^2\right\}\Bigg]\psi_1 = 0, \qquad \ldots\ldots(13) \end{aligned}$$

in which we have written

$${\omega_0}^2 = \frac{e\rho}{\epsilon m} = V\frac{d^2V}{dy^2} + \left(\frac{dV}{dy}\right)^2. \qquad \ldots\ldots(14)$$

(ii) *Linear Velocity Variation*

We shall not make any attempt to find a general solution of (13) for ψ_1 but shall consider only a linear increase of velocity with y and put

$$V = V_0 + Qy. \qquad \ldots\ldots(15)$$

Then from (1) $\rho = \epsilon m Q^2/e$. Therefore from (14)

$$Q = \omega_0, \qquad \ldots\ldots(16)$$

where ω_0 is the plasma resonance frequency.

On substituting

$$x = \frac{\omega}{\omega_0}\left(1 - \frac{hV}{\omega}\right) \qquad \text{......(17)}$$

into (13), the equation for ψ_1 reduces to

$$\frac{d^2\psi_1}{dx^2} + \frac{2x}{x^2-1}\frac{d\psi_1}{dx} - \psi_1 = 0. \qquad \text{......(18)}$$

Having found the equation that determines the motion of the electrons and the field inside the stream it is now necessary to ensure continuity of the field at the boundary of the stream. The next step is therefore to derive an expression for the radial admittance at the boundary of the perturbed stream. We use a method originally described by Hahn (1939).

If the boundaries of the unperturbed stream are $y = \pm a$ the boundaries of the perturbed stream oscillate about $y = \pm a$. For small perturbations it is sufficient to evaluate the magnetic field at the perturbed boundary $y = a + \Delta$ but to take this magnetic field as applying at the boundary $y = a$. In effect the perturbed boundary is replaced by the unperturbed boundary and the current between them is replaced by a surface current.

From Maxwell's equations we get an expression for H_x. Thus

$$\frac{\partial H_x}{\partial z} = \epsilon\frac{\partial E_y}{\partial t} - J_y. \qquad \text{......(19)}$$

Since

$$H_x = H_0 + H_1 e^{i\theta} \qquad \text{......(20)}$$

and $E_y = -\partial\phi/\partial y$, it follows from (19) that

$$H_1 = \frac{\omega\epsilon}{h}\frac{\partial\phi_1}{\partial y} - \frac{i}{h}j_y, \qquad \text{......(21)}$$

where $J_y = j_y e^{i\theta}$.

If Δ is the vertical displacement of the boundary we have, neglecting the displacement current,

$$H_x(a+\Delta) - H_x(a) = J_z\Delta = \rho V_z \Delta. \qquad \text{......(22)}$$

Now, as shown in Appendix II,

$$\Delta = \frac{1}{m}\frac{\partial\psi_1}{\partial y}\bigg/ i(\omega - Vh). \qquad \text{......(23)}$$

Hence on substituting from (21) and (23) into (22) we get

$$\begin{aligned} H_1(a+\Delta) &= \frac{\omega\epsilon}{h}\frac{\partial\phi_1}{\partial y} - \frac{i}{h}j_y - \frac{i\rho V}{m}\frac{\partial\psi_1}{\partial y}\bigg/(\omega - hV) \\ &= \frac{\omega\epsilon}{h}\frac{\partial\phi_1}{\partial y} - \frac{i\rho}{mh}\frac{1}{1-(hV/\omega)}\frac{\partial\psi_1}{\partial y}. \qquad \text{......(24)} \end{aligned}$$

The second line of (24) follows from (10).

Also $E_z = -\partial\phi/\partial z = E_1 e^{i\theta}$. Therefore

$$E_1 = ih\phi_1 = -\frac{\omega h}{e}\left(1 - \frac{hV}{\omega}\right)\psi_1. \qquad \text{......(25)}$$

Finally, the wave admittance in the positive direction of y is given by

$$Y = \frac{H_1}{E_1} = -\frac{i\omega\epsilon}{h^2}\frac{d}{dy}\log\phi_1 + \frac{i\omega\epsilon}{h^2}\left(\frac{\omega_0}{\omega}\right)^2\frac{1}{(1-hV/\omega)^2}\frac{d}{dy}\log\psi_1. \qquad \text{......(26)}$$

If we normalize this admittance by substituting

$$P = -\frac{ihY}{\epsilon\omega} \qquad \ldots\ldots(27)$$

and change the variable to x using (17), (26) becomes

$$P(x) = \frac{1}{x} + \left(1 - \frac{1}{x^2}\right)\frac{1}{\psi_1}\frac{d\psi_1}{dx}. \qquad \ldots\ldots(28)$$

We also require the value of the admittance of the field outside the stream at the stream boundaries. In the first instance we shall take the beam to be isolated in free space. The field outside the beam must then fall off in amplitude with distance from the beam. It is sufficient to represent this field by a simple evanescent TM-wave with phase velocity ω/h in the z direction. The admittance of such a wave is

$$Y = i\sqrt{\left(\frac{\epsilon}{\mu}\right)}\frac{k}{\sqrt{(h^2-k^2)}}.$$

In the quasi-stationary approximation which we are using, $h \gg k$, so we take

$$Y = i\sqrt{\left(\frac{\epsilon}{\mu}\right)}\frac{k}{h}. \qquad \ldots\ldots(29)$$

Normalizing it as in (27) gives

$$P(a) = 1. \qquad \ldots\ldots(30)$$

Similarly the admittance at $y = -a$ is

$$P(-a) = -1. \qquad \ldots\ldots(31)$$

(iii) *Determination of the Propagation Constant for the Free Stream*

The propagation constant h is determined by matching the admittance P at the boundaries of the stream.

Suppose we have found two independent solutions of the equation (18) for ψ_1. Denote them by $F(x)$ and $G(x)$. Then the general expression for ψ_1 is

$$\psi_1(x) = c_1\{F(x) + cG(x)\}. \qquad \ldots\ldots(32)$$

From (27) the admittance is

$$P(x) = \frac{1}{x} + \left(1 - \frac{1}{x^2}\right)\frac{F'(x) + cG'(x)}{F(x) + cG(x)} \qquad \ldots\ldots(33)$$

and the boundary conditions are

$$P(x_a) = -P(x_{-a}) = 1, \qquad \ldots\ldots(34)$$

where, by (15) and (17),

$$\left.\begin{aligned} x_a &= \frac{\omega}{\omega_0} - \frac{hV_0}{\omega_0}\left(1 + \frac{a\omega_0}{V_0}\right) \\ \text{and} \qquad x_{-a} &= \frac{\omega}{\omega_0} - \frac{hV_0}{\omega_0}\left(1 - \frac{a\omega_0}{V_0}\right). \end{aligned}\right\} \qquad \ldots\ldots(35)$$

It will be seen from (35) that the unknown propagation constant and the frequency both appear in the expressions for the boundaries $x_{\pm a}$. This complicates the procedure for evaluating h.

Before we discuss the results of solving (33) and (34) we shall complete the analysis by considering a stream loaded with admittance sheets at the boundaries.

(iv) *Linear Velocity Variation with Loaded Guide*

In order to study the interplay of the electron space charge and a slow TM-waveguide wave we shall analyse the following system. The mean velocity of electrons increases from 0 at $y=-a$ to $2\omega_0 a$ at $y=+a$. At the boundary $y=-a$ we assume a perfectly conducting sheet and at $y=a$ an impedance sheet with such properties as are necessary for the propagation of a wave of phase velocity v_0 between $y=-a$ and $y=+a$ in the absence of free electrons.

For this electron-free wave

$$\left.\begin{aligned} E_z &= E_0 \sinh T(y+a), \\ H_x &= -\frac{i\omega\epsilon}{T} E_0 \cosh T(y+a), \\ Y(a) &= -\frac{i\omega\epsilon}{T} \coth(2aT), \end{aligned}\right\} \qquad \ldots\ldots(36)$$

where $T^2 = h_0{}^2 - k^2$ and $h_0 = \omega/v_0$.

For $v_0 \ll c$

$$Y(a) = -\frac{i\omega\epsilon}{h_0} \coth(2ah_0). \qquad \ldots\ldots(37)$$

The normalized admittance at $y=a$ is therefore

$$P(x_a) = -\frac{h}{h_0} \coth(2ah_0). \qquad \ldots\ldots(38)$$

Consequently from (37), taking the velocity of electrons at $y=-a$ to be zero so that $V_0 = \omega_0 a$,

$$P(x_{-a}) = \infty$$

and

$$P(x_a) = -\frac{h}{h_0} \coth\left(\frac{2h_0 V_0}{\omega_0}\right).$$

Let η be the ratio of the phase velocity of the wave in the empty guide to the mean velocity of the fastest electrons. Then

$$\eta = \frac{v_0}{2V_0} = x_{-a} \frac{\omega_0}{2h_0 V_0}, \qquad \ldots\ldots(39)$$

and from (37)

$$h = \eta h_0 \left(1 - \frac{x_a}{x_{-a}}\right), \qquad \ldots\ldots(40)$$

where

$$x_{-a} = \frac{\omega}{\omega_0}.$$

The boundary conditions then become

$$\left.\begin{aligned} P(x_{-a}) &= \infty, \\ P(x_a) &= -\eta \left(1 - \frac{x_a}{x_{-a}}\right) \coth\left(\frac{x_{-a}}{\eta}\right), \end{aligned}\right\} \qquad \ldots\ldots(41)$$

where

$$P(x) = \frac{1}{x} + \left(1 - \frac{1}{x^2}\right) \frac{F'(x) + cG'(x)}{F(x) + cG(x)}.$$

From the first condition of (41) we deduce that

$$c = -\frac{F(x_{-a})}{G(x_{-a})}. \qquad \ldots\ldots(42)$$

§3. THE RESULTS

The analysis presented in §2 is based on a quasi-stationary approximation of which the main assumptions are (*a*) that the phase velocity of the propagated wave is much less than *c*, (*b*) that the effect on the electrons of their own magnetic field is negligible in comparison with the effect of the externally applied magnetic field, (*c*) that the oscillating field is of small amplitude compared with the D.C. field. In the results to be described in this section attention has been focused on only one mode of wave propagation. It is the lowest-order mode, which in most applications would be the dominant mode. There is, however, an infinite set of higher-order modes, some of which may travel against the stream and some may be amplifying. We shall not pursue their study further in this work.

(i) *Solutions of the Equation for the Action Function*

The key to the behaviour of the slipping stream as an amplifier is contained in the functions $F(x)$ and $G(x)$, two independent solutions of the wave equation for ψ_1, equation (18). Since x depends on h, which is in general complex, we must study F and G for complex values of the argument. We have therefore tabulated $F(x)$ and $G(x)$ defined in the following way:

$$F(1,0)=1, \qquad F'(1,0)=0,$$

$$G(x)=c_2[F(x)\log c_1(1-x)+g(x)], \qquad g(1,0)=0,$$

$$G(x)=F(-x), \qquad G'(x)=-F'(-x).$$

Then $F(x)=c_2[G(x)\log c_1(1+x)+g(-x)]$. The property of these functions which is of primary interest to us in this discussion is that $F(x)$ and $G(x)$ have logarithmic singularities at the points $(-1,0)$ and $(1,0)$ respectively. If we refer to the definition of x at equation (17) we find that at the singularities,

$$\omega=\frac{\pm\,\omega_0}{1-hV/\omega}.$$

But the right-hand side of this equation is the plasma frequency ω_0 increased by the Doppler factor due to its motion with velocity V in a medium in which waves travel with velocity ω/h. The singularity therefore corresponds to a condition of resonance between the applied frequency ω and the apparent plasma frequency measured by a stationary observer.

(ii) *Properties of the Lowest-order Modes in a Free Slipping Stream*

In the analysis and discussion of the results it is convenient to describe the velocity range of the electrons in terms of $\alpha=(V_a-V_{-a})/2V_0$. When $\alpha=1$ the velocity is zero at one edge of the beam and $2V_0$ at the other edge.

In Figure 1 we show the imaginary part of the propagation constant h as a function of frequency for three values of α. These results were obtained by solving the equations (33) and (34) of §2 (iii). The frequency axis is marked in units of the plasma resonance frequency ω_0 and h is measured in units of ω_0/V_0.

Consider firstly the spectrum for $\alpha=0{\cdot}25$. Figure 1 shows apparently three amplifying and three attenuating waves, taking the minus and plus signs of $\operatorname{Im}(hV_0/\omega_0)$ respectively. (In the following discussion we shall confine our attention to the amplifying waves only.) The wave marked (1) has an amplification range $0<\omega<\omega_1$, and a dependence of gain on frequency which bears a strong resemblance to that of a two-beam tube. Comparison of the two shows that the gain–frequency characteristic of the slipping-stream tube is very nearly the same as that of a two-beam tube which has stream velocities V_a and V_{-a} but plasma resonance frequency $\omega_0/2$. Since the Larmor frequency ω_L is equal to $\omega_0/2$, it appears that the plasma resonance frequency of the two-beam tube, whose electron space-charge field is balanced by an electric field of positive charges, should be compared with the Larmor frequency of the slipping-stream tube, whose space-charge field is balanced by a transverse magnetic field. As in the two-beam tube the phase velocity is about equal to the mean velocity V_0 of the electrons. Thus wave (1) amplifies by the same mechanism as the dominant amplifying wave of the two-beam tube. It is therefore convenient to call it the two-beam tube wave.

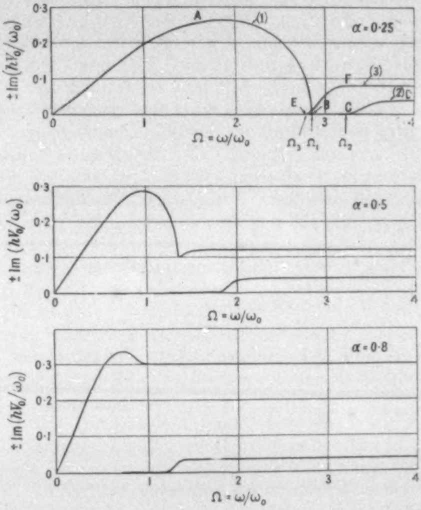


Figure 1. Dependence of rate of gain on frequency in a free slipping stream.

Waves (2) and (3), however, have no counterpart in the two-beam tube. On the contrary they are of the same type as the amplifying wave in a travelling-wave tube. The effective reactive impedance sheet, which guides the slow wave, is provided by a resonance layer within the stream itself. In order to appreciate this phenomenon the reader is referred to Figures 2 to 5.

In Figure 2 the loci of the points x_a and x_{-a}, defined by (35), are plotted for a range of values of α. They are values of the variable x, which varies linearly across the beam from x_{-a} at the low-velocity edge to x_a at the high-velocity edge. In order to assist in the explanation by the following argument a sketch of the loci of x_a and x_{-a} for $\alpha \simeq 0{\cdot}25$ is given in Figure 3.

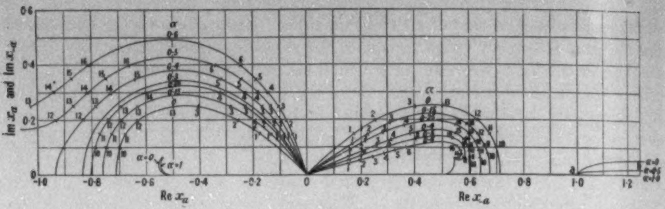
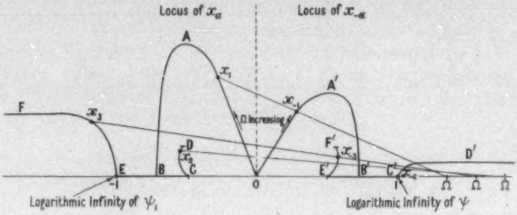
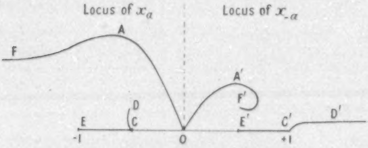


Figure 2. The loci of the points x_a and x_{-a} for a range of values of $\alpha=(V_a-V_{-a})/2V_0$. The numbers mark corresponding points of x_a and x_{-a}.



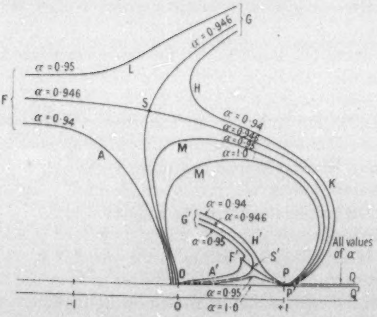
Locus of x_a
Locus of x_{-a}
A
x_1
A′
x_{-1}
F
x_3
Ω increasing
D
x_2
C
F′
x_{-3}
D′
E
B
-1
0
E′
B′
C′
1
x_{-2}
Ω
Ω
Ω
Logarithmic Infinity of ψ_1
Logarithmic Infinity of ψ

Figure 3. Sketch of the loci of x_a and x_{-a} for $\alpha<0{\cdot}42$.



Locus of x_a
Locus of x_{-a}
A
F
A′
F′
D
C
E
-1
0
E′
C′
+1
D′

Figure 4. Sketch of the loci of x_a and x_{-a} for $0{\cdot}42<\alpha<0{\cdot}9$.



$\alpha=0{\cdot}946$
G
$\alpha=0{\cdot}95$
L
F
$\alpha=0{\cdot}946$
S
H
$\alpha=0{\cdot}94$
A
$\alpha=0{\cdot}94$
$\alpha=0{\cdot}946$
$\alpha=0{\cdot}95$
$\alpha=1{\cdot}0$
M
M
K
$\alpha=0{\cdot}94$
G′
$\alpha=0{\cdot}946$
$\alpha=0{\cdot}95$
H′
All values of α
F′
S′
O
A′
P
Q
$\alpha=0{\cdot}95$
P′
Q′
-1
0
$\alpha=1{\cdot}0$
+1

Figure 5 Sketch of the loci of x_a and x_{-a} for $0{\cdot}9<\alpha<1{\cdot}0$.

From the expression for x, namely

$$x=\frac{\omega}{\omega_0}-\frac{h}{\omega_0}(V_0+\omega_0 y),$$

it follows that the value of x at the point of intersection of the straight line joining x_a to x_{-a} and the real axis is the frequency $\Omega=\omega/\omega_0$ and that

$$\operatorname{Im}(hV_0/\omega_0)=\tfrac{1}{2}(\operatorname{Im}x_a+\operatorname{Im}x_{-a}).$$

The straight line joining x_a and x_{-a} also defines the distribution of the action function ψ_1 across the beam. Now, as indicated in § 3 (i) ψ_1 has logarithmic singularities at the points $(-1,0)$ and $(1,0)$. If, therefore, the line passes near to either of these points ψ_1, and consequently E_{1z}, will have a pronounced maximum at the level in the stream that corresponds to the nearest point of approach. Moreover it is known (Ramo 1939) from the study of wave propagation in a uniform electron stream of velocity V that a slow TM-wave can propagate along the stream provided $\operatorname{Re}(\omega-Vh)\simeq\pm\omega_0$. In terms of x this condition is $\operatorname{Re}x\simeq\pm 1$. It signifies the state of resonance between the applied frequency and the plasma resonance frequency as measured by a stationary observer. If the real parts of x_{-a} and x_a enclose the point $x=1$, or the point $x=-1$, the resonance condition for the propagation of a slow TM-wave will be approximately satisfied at one level in the stream, which can be appropriately called the resonance level. The resonance level behaves therefore like the corrugated surface or reactive impedance sheet of a travelling-wave tube: it can guide a slow wave. The axial electric field falls off rapidly away from the resonance level in just the same way as from the effective impedance sheet of a travelling-wave tube. If therefore a resonance layer occurs in the beam of a slipping-stream tube we should expect the wave supported by it to interact with electrons in other levels in the beam in much the same way as the wave guided by the helix or corrugated structure reacts with the electrons in a travelling-wave tube.

In the light of these remarks let us consider how the loci of x_a and x_{-a} move as the frequency is increased from zero. In Figure 3 x_a and x_{-a} lie at the origin when $\Omega=\omega/\omega_0=1$. As Ω is increased x_a moves out along the curve OAB and x_{-a} along OA′B′. For these waves the real parts of x_a and x_{-a} do not enclose either $x=1$ or $x=-1$ and so no resonance layer occurs in the beam. This gives the two-beam tube wave of Figure 1. However, when x_a reaches the point B its path forks. One branch moves along the real axis to C then up to the limiting point D, where $x=-0{\cdot}542+0{\cdot}061i$, while x_{-a} moves to the resonance point C′ and then out along C′D′. In the range BC the propagation constant is entirely real but it acquires an imaginary part beyond C, as shown for $\alpha=0{\cdot}25$ in Figure 1. The cut-off frequency at C is simply related to α by the formula

$$\Omega_2=(0{\cdot}754/\alpha)+0{\cdot}247.$$

In the range C to D it is clear from Figure 3 that a resonance layer occurs in the stream. It appears at the low-velocity edge of the stream at frequency Ω_2 and moves towards the high-velocity edge as the frequency is increased. The phase velocity also increases, as shown by the curve CD in Figure 6, in keeping with the velocity of the electrons at the resonance layer. As Ω tends to infinity $\operatorname{Im}(hV_0/\omega_0)$ tends to $0{\cdot}061/(1+\alpha)$, which gives a gain of $0{\cdot}53\omega_0/V_a$ decibels per unit length.

If we now return to the point B and follow the other branch we find that x_a moves along the real axis to the singularity at E and then along EF, while x_{-a} moves to E′, where $x=0{\cdot}507$, and up to F′, where $x=0{\cdot}542+0{\cdot}061i$. Figure 1 shows that as x_a moves from B to E the frequency decreases to $\Omega_3=(0{\cdot}754/\alpha)-0{\cdot}247$ and $\operatorname{Im} h=0$, but as x_a moves out along EF the frequency again increases and $\operatorname{Im}(hV_0/\omega_0)$ increases asymptotically to the value $0{\cdot}061/(1-\alpha)$ as $\Omega\to\infty$. This gives a gain of $0{\cdot}53\omega_0/V_{-a}$ decibels per unit length. It is evident from Figure 3 that in the range EF a resonance layer occurs in the beam and that it moves from the high-velocity edge towards the low-velocity edge as the frequency increases from Ω_3. The phase velocity also decreases, as shown by the curve EF in Figure 6.

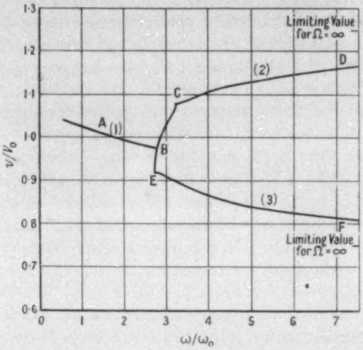


Figure 6. The phase velocity as a function of frequency for $\alpha=0{\cdot}25$.

We can now explain why wave (3) has a higher gain than wave (2) for $\Omega\gg\Omega_2$. It is because the gain for a travelling-wave tube is larger the lower the phase velocity of the wave guided by its effective impedance sheet. In the slipping-stream tube this wave has phase velocity about equal to the velocity of the electrons at the resonance layer and for $\Omega\gg\Omega_2$ the resonance layer is near the low-velocity edge with wave (3) and near the high-velocity edge with wave (2). Therefore the gain of wave (3) should be higher than the gain of wave (2).

The gains of waves (2) and (3) do not increase with frequency according to the law $\operatorname{Im}(hV_0/\omega_0)\propto\omega^{1/3}$ of the travelling-wave tube. That they should not do so is apparent when it is remembered that the slow wave guided by the resonance layer will gather energy only from electrons which have velocities slightly greater than the phase velocity of the wave. Slower electrons will absorb energy from the wave. In a travelling-wave tube all the electrons travel slightly faster than the wave.

Having discussed the characteristics of the waves for $\alpha\simeq0{\cdot}25$ in some detail let us now discuss the effect of changing α. Referring to Figure 2 we see that as α is increased from zero to 0·42 the branch point B of Figure 3 moves from $x=-0{\cdot}715$ to $x=-1$. When $\alpha>0{\cdot}42$ the locus of x_a splits into two quite separate parts as shown in Figure 4. Waves (1) and (3) merge into one without any overlap in the gain–frequency characteristic of Figure 1. The examples $\alpha=0{\cdot}5$

and $\alpha=0{\cdot}8$ of Figure 1 illustrate this effect. They also show the steep rise in gain at high frequencies as $\alpha\to 1$, according to the law $\mathrm{Im}\,(hV_0/\omega_0)\sim 0{\cdot}061/(1-\alpha)$. Wave (2) however, starts at $\Omega_3=(0{\cdot}754/\alpha)-0{\cdot}247$ when x_a is at E. Between Ω_2 and $\Omega_3=(0{\cdot}754/\alpha)+0{\cdot}247$ it has a real propagation constant and when $\Omega>\Omega_2$ it has a gain which tends to the limiting value $\mathrm{Im}\,(hV_0/\omega_0)\sim 0{\cdot}061/(1+\alpha)$ as $\Omega\to\infty$.

That the behaviour of x_a and x_{-a} indicated in Figure 4 is not the whole story becomes evident when we consider what happens as α approaches the value unity. It is found that the loci of x_a and x_{-a} change in the manner shown in Figure 5. It now appears that the locus of x_a has a third branch GHKPQ (for clarity we have omitted the branch ECD from Figure 5). Moreover there is a critical value of α for which branches OAF and GHKPQ meet. This new branch passes through the singularity at $x=+1$ and runs along the real axis from $x=1$ to $x=\infty$. When $\alpha>0{\cdot}946$, $\mathrm{Re}\,x_a>\mathrm{Re}\,x_{-a}$ along most of the curve GHKP; then it gives reverse waves, of which only the attenuating ones given by the negative sign of $\mathrm{Im}\,h$ are of practical importance. Amplifying reverse waves are also given by this theory but they must always be accompanied by attenuating waves of larger amplitude.* When $\alpha=0{\cdot}946$ branches OAF and GHK meet at the point S, where $x_a=-0{\cdot}30+1{\cdot}98i$, and when $\alpha>0{\cdot}946$ they split into an upper branch FLG and a lower branch OMP. The corresponding loci of x_{-a} are indicated by dashed letters in Figure 5.

The most interesting feature of Figure 5 is the continued rise of the curve LF as $\alpha\to 1$. It can be shown that the curve LF tends to the asymptote

$$\mathrm{Im}\,x_a=0{\cdot}061(1+\alpha)/(1-\alpha) \quad \text{as} \quad x_{-a}\to 0{\cdot}542+0{\cdot}061i$$

and this gives the high gain, $\mathrm{Im}\,(hV_0/\omega_0)=0{\cdot}061/(1-\alpha)$, as $\alpha\to 1$ at frequencies above ω_0. It should not surprise one that the gain tends to infinity on this theory as $\alpha\to 1$. As the limit is approached both the velocity of the electrons at one edge of the beam and the velocity of the slow wave guided by the resonance layer tend to zero. The system should therefore behave as a travelling-wave tube with vanishing electron velocity which, in theory, would also have infinite gain.

It is worth while at this stage to compare the theoretical expressions for the maximum gain of the travelling-wave tube (TWT), two-beam tube (TBT) and slipping-stream tube (SST). We obtain the following values, expressed in decibels per unit length:

$$G_{\mathrm{TWT}}\simeq 6\left(\frac{\omega}{\omega_0}\right)^{1/3}\frac{\omega_0}{V},$$

$G_{\mathrm{TBT}}\simeq 4{\cdot}3\,\dfrac{\omega_0}{\overline{V}}$, where $\overline{V}$ is the mean velocity of the two electron streams,

$G_{\mathrm{SST}}\simeq 2{\cdot}1\,\dfrac{\omega_0}{\overline{V}}$ for the TBT-mode

and

$G_{\mathrm{SST}}\simeq 0{\cdot}53\,\dfrac{\omega_0}{V_{-a}}$ for the TWT-mode.

---

* A discussion of reverse waves will be found in a forthcoming paper by G. G. Macfarlane and Mrs. A. M. Woodward.

These expressions apply to ideal systems in which no metal and other losses occur and in which the electrons move in regular streams without any random thermal motions.

Maximum amplification is obtained in all cases when the electron velocity is a minimum. It is therefore reasonable to compare the systems for the same minimum velocity. Moreover, in order to obtain gain at frequencies above ω_0, with the two-beam tube it is necessary for V_{-a} to be nearly equal to $\overline{V}$. Therefore for comparison of the performance of these ideal tubes as amplifiers for frequencies much higher than ω_0 we can put $V=\overline{V}=V_{-a}$. Then we get

$$G_{\mathrm{TWT}}:G_{\mathrm{TBT}}:G_{\mathrm{SST}}(\text{TBT-mode}):G_{\mathrm{SST}}(\text{TWT-mode}) = 6(\omega/\omega_0)^{1/3}:4{\cdot}3:2{\cdot}1:0{\cdot}53.$$

A more detailed discussion of the effects of losses, thermal motion of the electrons, and possible feeding systems is beyond the scope of this paper.

It still remains to mention another result which may be important in consideration of the use to which a slipping-stream tube might be put. In Figure 5 it is shown that the portion of the real axis from $x=1$ to $x=\infty$ is a common locus for x_a and x_{-a} for all values of α. There is no gain or attenuation for these waves and they are reverse uniform waves. When α is equal to unity the phase velocity depends on frequency for $\Omega>1$ as shown in Figure 7. These reverse waves may constitute a serious drawback to the use of a slipping-stream tube as an amplifier unless it can be found how to avoid exciting them. The fact, however, that Haeff has obtained a stable amplifier using a slipping stream is encouraging.

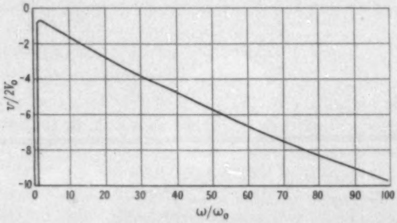


Figure 7. The phase velocity of the reverse uniform waves as a function of frequency, $\alpha=1$.

Summing up the discussion we can say that a slipping-stream tube will behave differently according to the fractional velocity range of the electrons: if

$$(V_a-V_{-a})/(V_a+V_{-a})=\alpha<0{\cdot}42$$

it behaves primarily like a two-beam tube with cut-off frequency $\Omega_1=1/\alpha\sqrt{2}$ but it has only about one half its maximum gain, viz. $2{\cdot}1\omega_0/V_0$ decibels per unit length. In addition it has a gain of $0{\cdot}53\omega_0/V_{-a}$ for $\Omega\gg\Omega_1$. However, if $\alpha>0{\cdot}42$ it behaves primarily like a two-wave tube with a gain of $0{\cdot}53\omega_0/V_{-a}$ decibels per unit length. Comparison with the travelling-wave tube and two-beam tube on the basis of the same minimum electron velocity shows that the slipping-stream tube has a lower maximum rate of gain than either of the other two.

(iii) *Properties of the Amplifying Wave in a Slipping-stream Travelling-wave Tube*

In this section we give results for a slipping stream inside a waveguide. The waveguide is such that a slow TM-wave of phase velocity v_0 can travel along it when there is no electron beam inside it. The velocity is taken to vary from 0 to $2V_0$ ($\alpha=1$). The system is a form of travelling-wave tube using a slipping stream of electrons instead of a uniform stream.

The theory, by means of which the following results were obtained, is set out in §2(iv). It was shown in that section that the waveguide could be defined in terms of an impedance sheet. This concept obviates the necessity of specifying the physical form of loaded guide required to support a slow wave of TM-type. It effects at once a unification and simplification because a TM-wave of phase velocity v_0 with the field distribution given in §2(iv) can be guided by different physical structures. The object of our analysis is to discover how the propagation constant h varies with frequency for a range of values of v_0 in the neighbourhood of $2V_0$. In most physical structures used to obtain such slow waves the phase velocity v_0 will vary with frequency. By giving a map of h marked in contours of constant $2V_0/v_0$ and ω/ω_0 the dependence of the propagation constant on frequency can be quickly determined when the variation of v_0 with frequency is known. Such a contour map is shown in Figure 8, in which, instead of h, the quantity given is $x_a=(\omega/\omega_0)-(2hV_0/\omega_0)$. Since $\alpha=1$, $x_{-a}=\omega/\omega_0$, so that the real component of h can be calculated immediately from the curves. The ordinate is simply $\mathrm{Im}\,(2hV_0/\omega_0)$. From the figure it is therefore clear that the maximum rate of gain is obtained when $v_0 \simeq 2V_0$, that is when the phase velocity in the empty guide is about equal to the velocity of the fastest electrons. It is interesting to compare these results with the corresponding results for a travelling-wave tube with uniform beam, which are shown in Figure 9 (Woodward, unpublished work).

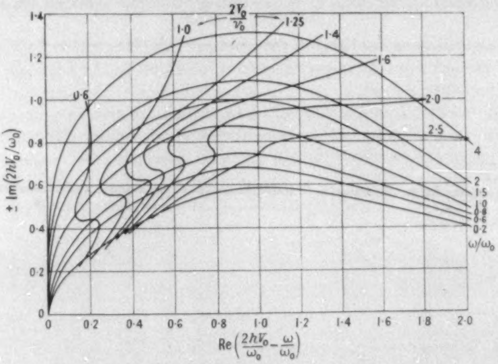


Figure 8. Contour map of the propagation constant for the slipping stream travelling-wave tube v_0 is the phase velocity of the TM-wave in the empty guide.

In this figure the velocity of the electrons is V and the phase velocity in the empty guide is v. The similarity of the characteristics at any rate for $\omega>\omega_0$ is immediately evident. The maximum gain is slightly greater for the slipping-stream tube. There is, however, one noticeable difference. When $v_0<V$ there is an

upper limit to the gain with the travelling-wave tube, but when $v_0<2V_0$ there is no upper limit with the loaded slipping-stream tube. The reason for this is that the electron velocity ranges in value from 0 to $2V_0$ in the slipping-stream tube so that there are always some electrons for which the condition, $v_0 \geqslant V$, for gain at the highest frequencies is satisfied. However, for $v_0>2V_0$ the gain falls to low values at high frequencies because the electrons which can give up energy to the waves are further from the impedance sheet. Moreover the field falls off rapidly from the impedance sheet.

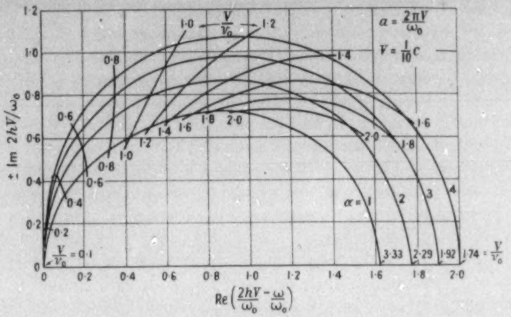


Figure 9. Contour map of the propagation constant for a uniform stream travelling-wave tube.

In Figure 10 we have plotted the phase velocity of the wave as a function of frequency. It is always less than the velocity of the fastest electrons.

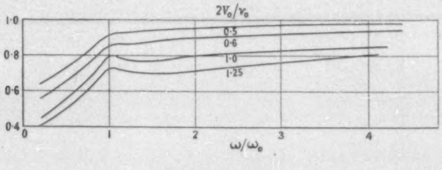


Figure 10. The phase velocity v as a function of frequency for the slipping stream travelling-wave tube. The parameter is $2V_0/v_0$, where v_0 is the phase velocity in the empty guide.

§ 4. ADDITIONAL NOTES

(*a*) The non-oscillatory slipping stream requires applied electric and magnetic fields to ensure parallel motion of the electrons. From (3) the magnetic flux density is

$$B_x=\frac{m}{e}\omega_0=5{\cdot}68\times10^{-8}\omega_0 \text{ gauss}.$$

From (1) the potential in the beam is $\phi=mV^2/2e$. If the electric field is maintained by a potential difference between two parallel plates at $y=\pm b$ the potential on the plates must be

$$\phi(b)=\frac{m}{2e}V_2{}^2+(b-a)\frac{m\omega_0}{e}V_2; \qquad \phi(-b)=\frac{m}{2e}V_1{}^2-(b-a)\frac{m\omega_0}{e}V_1,$$

where $\phi=0$ is the potential of the cathode from which the electrons are emitted (*b*) In order to study the basic phenomena we have described a planar form of slipping-stream tube. Other forms can be thought of which may be of greater practicability. For example an annular beam might be used between the inner and outer conductors of a coaxial line; the magnetic field, which would be azimuthal, could be produced by axial currents in the inner and outer conductors.

(*c*) From the small-signal theory it is not possible to deduce the behaviour of the system at large signal amplitudes. Nevertheless it is interesting to try to discuss a few general points connected with the transformation of energy in the system.

It can be easily shown that the mean flow of energy across the section of the tube is

$$W = -\int_{\Gamma} \overline{(\mathbf{E}\wedge\mathbf{H})} \,.\, \mathbf{n}\, da - \frac{m}{2e}\int \overline{(\rho V^2)\mathbf{V}} \,.\, \mathbf{n}\, da. \qquad \ldots\ldots(43)$$

This latter integral is the mean rate at which kinetic energy of the electrons is passing the section Γ.

At the input section the R.F. energy can be neglected, $V = V_0 + \omega_0 y$, and Γ extends from $y = -a$ to $y = +a$. Evaluating the integral we get

$$W = \frac{m}{e}\rho a(V_0^2 + a^2\omega_0^2)V_0. \qquad \ldots\ldots(44)$$

At a section farther along the tube this power is partly converted into oscillatory power. If we still think of the beam in two parts, one D.C. and the other A.C., we see that the A.C. power can increase at the expense of the D.C. power either by a decrease in mean velocity or a decrease of density of the electrons or by both. Now in a travelling-wave tube with strong axial magnetic field the D.C. power comes primarily from a decrease in mean velocity because lateral spreading is inhibited, but in the slipping-stream tube the beam is at liberty to spread and lose average density as well as mean velocity. On this account it seems likely that there will be a more efficient transfer of power from D.C. to A.C. at large amplitude with the slipping-stream tube.

ACKNOWLEDGMENT

The work described above has been carried out for the Department of Scientific and Industrial Research and is published by permission of the Director of Radio Research, D.S.I.R., the Chief Scientist, Ministry of Supply, and the Controller of H.M. Stationery Office.

REFERENCES

BUNEMANN, O., 1944, *C. V. D. Report Mag.*, **37.** A useful account of this work is given by L. R. Walker in *Microwave Magnetrons*, No. 6 of *Radiation Laboratory Series* (McGraw-Hill Book Co.), 1948, pp. 253–264.

GABOR, D., 1945, *Proc. Inst. Radio Engrs.*, *N.Y.*, **33,** 804.

HAEFF, A. V., 1949, *Proc. Inst. Radio Engrs.*, *N.Y.*, **37,** 4.

HAHN, W. C., 1939, *Gen. Elect. Rev.*, **42,** 258.

RAMO, S., 1939, *Phys. Rev.*, **56,** 276.

RAMSEY, A. S., 1935, *A Treatise on Hydrodynamics*, Pt. II, 4th edn. (London : G. Bell and Son), p. 13.

SMITH, L. P., and HARTMAN, P. L., 1940, *J. Appl. Phys.*, **11,** 220.

APPENDIX I

DERIVATION OF THE HAMILTON–JACOBI EQUATION

The following discussion is restricted to non-relativistic motion.

In terms of the scalar potential ϕ and the vector potential $\mathbf{A}$ the equation of motion of an electron in Newtonian form is

$$m\frac{d\mathbf{V}}{dt}=-e\mathbf{E}-e\mathbf{V}\wedge\mathbf{B}=e\left(\frac{\partial\mathbf{A}}{\partial t}+\operatorname{grad}\phi-\mathbf{V}\wedge\operatorname{curl}\mathbf{A}\right). \qquad \ldots\ldots(45)$$

For electrons moving in a stream the left hand side of (45) can be transformed, as in classical hydrodynamics, to give

$$\frac{d\mathbf{V}}{dt}=\frac{\partial\mathbf{V}}{\partial t}+(\mathbf{V}\,.\operatorname{grad})\mathbf{V}=\frac{\partial\mathbf{V}}{\partial t}+\tfrac{1}{2}\operatorname{grad}V^2-\mathbf{V}\wedge\operatorname{curl}\mathbf{V}. \qquad \ldots\ldots(46)$$

Then (45) becomes

$$\frac{\partial}{\partial t}(m\mathbf{V}-e\mathbf{A})=\mathbf{V}\wedge\operatorname{curl}(m\mathbf{V}-e\mathbf{A})+\operatorname{grad}(e\phi-\tfrac{1}{2}mV^2). \qquad \ldots\ldots(47)$$

If the motion is such that

$$\mathbf{V}\wedge\operatorname{curl}(m\mathbf{V}-e\mathbf{A})=0, \qquad \ldots\ldots(48)$$

$m\mathbf{V}-e\mathbf{A}$ must be the gradient of some scalar, ψ say. Then

$$m\mathbf{V}-e\mathbf{A}=\operatorname{grad}\psi \qquad \ldots\ldots(49)$$

and, on integrating (47),

$$\frac{\partial\psi}{\partial t}+\tfrac{1}{2}mV^2-e\phi=0, \qquad \ldots\ldots(50)$$

since the constant of integration, which may be a function of time, can always be reduced to zero by suitable choice of ψ.

Eliminating $\mathbf{V}$ from (49) and (50) gives

$$\frac{\partial\psi}{\partial t}+\tfrac{1}{2}m(e\mathbf{A}+\operatorname{grad}\psi)^2=e\phi,$$

which is the Hamilton–Jacobi equation in § 2 (i).

It can be shown (Gabor 1945) that the condition $\operatorname{curl} m\mathbf{V}=e\mathbf{B}$ is satisfied at all points of the motion if the electrons leave the cathode with zero velocity and there is no component of magnetic field normal to the cathode surface.

Given that $\operatorname{curl} m\mathbf{V}=e\mathbf{B}$, the two-dimensional problem treated in this paper demands that the magnetic field maintaining the velocity distributions be at right angles to the plane of the motion, and, in addition, be of a specified magnitude. For if $\mathbf{V}=[0, V_y, V_z]$ and $\mathbf{V}$ is independent of x, then

$$\operatorname{curl}\mathbf{V}=\left[\frac{\partial V_z}{\partial y}-\frac{\partial V_y}{\partial z},0,0\right],$$

so that $\mathbf{B}=[B_x, 0, 0]$, where

$$B_x=\frac{m}{e}\left(\frac{\partial V_z}{\partial y}-\frac{\partial V_y}{\partial z}\right).$$

APPENDIX II

If the equation of the boundary surface is

$$F(y, z, t) = y - a - \Delta e^{i\omega t - ihz} = 0,$$

the function F must satisfy the equation (Ramsey 1935)

$$\frac{\partial E}{\partial t} + V_y \frac{\partial F}{\partial y} + V_z \frac{\partial F}{\partial z} = 0.$$

Since $V_z = V$ and $V_y = \frac{1}{m}\frac{\partial \psi_1}{\partial y} e^{i\omega t - ihz}$, this gives

$$\Delta = \frac{1}{m}\frac{\partial \psi_1}{\partial y} \Big/ i(\omega - Vh).$$

Theory of the Production of an Ionized Layer in a Non-Isothermal Atmosphere Neglecting the Earth's Curvature, and its Application to Experimental Results

By J. A. GLEDHILL and M. E. SZENDREI
Rhodes University College, Grahamstown, South Africa

Communicated by M. Blackman; MS. received 31*st March* 1949, *and in amended form* 26*th January* 1950

ABSTRACT. A new theory of ionospheric layer formation is developed, in which the temperature is assumed to vary linearly with height. The equations are compared at each step with those obtained by Chapman in his theory of layer formation in an isothermal atmosphere. The equations for the maximum of electron density and its height are also given. The effect of the parameters on the shape of the layer is shown in graphical form.

The equations are somewhat complex in form, but an ingenious graphical method has been devised suitable for the application of the theory to results given in the form of a Booker and Seaton parabolic distribution of electronic density with height. By application of the theory to mean hourly values for four summer months in South Africa, values are obtained for the temperature gradient, the temperature at 200 km., and its variation over the middle part of the day. The results obtained are in accordance with previous estimates and offer numerical confirmation of the theory that the atmosphere expands bodily upwards during the middle part of a summer day.

§ 1. THE THEORY

1 (i). *Introduction*

THE most prominent among the theories put forward to account for the formation of ionospheric layers is that of Chapman (1931, 1939), which attributes their production to the ionizing action of solar ultra-violet radiation on the earth's atmosphere. Other possible causes, such as corpuscular radiation, cosmic rays, alpha and beta rays from various sources, meteors, etc., are entirely neglected; owing to the irregular nature of most of them, they are intractable to theoretical treatment at present. In Chapman's theory a large number of simplifying assumptions is necessarily made:

(i) the radiation is monochromatic;
(ii) the density of the atmosphere, on which the radiation is incident, varies exponentially with height;
(iii) the atmosphere is at rest and uniform in composition;
(iv) the temperature is the same at all points in the atmosphere at all times;
(v) the force of gravitation is constant with height;
(vi) all the absorbed energy causes ionization.

Charts showing the ionization of the upper atmosphere have been constructed by Millington (1932, 1935), using Chapman's theory, but do not seem to have had much application.

Various approximate theories have been put forward by Hulburt (1928, 1938); finally an attempt was made (Hulburt 1939a) to fit a theoretical curve to the observed mean values of maximum electron density for the E layer.

Comparison shows that Chapman's and Hulburt's theories are the same in fundamental content, being based on the same initial assumptions. Their application to the E region has met with considerable success, but few attempts have been made to apply them to the higher more irregular regions.

To develop a theory which will represent more nearly existing conditions in the F regions, it seems advisable to reject Chapman's assumption of an isothermal atmosphere and, instead, to allow the temperature to vary with height in a known way; as the simplest variation possible, it has been assumed in this paper that the temperature varies linearly with height from a given level upwards. The distribution of molecules with height in such an atmosphere will no longer be exponential and will first be derived.

1 (ii). *Distribution of Molecules with Height*

Assuming no vertical movement of the atmosphere, the increase in pressure dp corresponding to an increase in height dh is given by

$$-dp = \mathbf{g}\rho\, dh, \qquad \ldots\ldots(1)$$

where $\mathbf{g}$ is the acceleration due to gravity and ρ the density at height h. We use the simple gas laws, and put n' for the number of molecules per cm^3 at height h and n_0' for the corresponding value at a reference level h_0, the temperature at this height being T_0. On substitution and integration between the limits n_0' and n', corresponding to heights h_0 and h, we have

$$n' = n_0' \frac{T_0}{T} \exp\left(-\frac{m\mathbf{g}}{\boldsymbol{k}} \int_{h_0}^{h} \frac{dh}{T} \right), \qquad \ldots\ldots(2)$$

where $\boldsymbol{k}$ = Boltzmann's constant, and m = mean molecular mass.

In this paper we limit the discussion to the case where T is a linear function of height and is given by

$$T = T_0 + \gamma(h - h_0). \qquad \ldots\ldots(3)$$

Substitution in equation (2) and integration then yields

$$n' = n_0' \left\{ 1 + \gamma \frac{h - h_0}{T_0} \right\}^{-(1 + C/\gamma)}, \qquad \ldots\ldots(4)$$

where we have written

$$C=\frac{m\boldsymbol{g}}{\boldsymbol{k}}. \qquad \text{......(5)}$$

On expressing this as a binomial expansion and inserting the limit $\gamma=0$, equation (4) reduces to

$$n'=n_0'\exp\left\{-\frac{C}{T_0}(h-h_0)\right\}, \qquad \text{......(6)}$$

which is the exponential distribution assumed by Chapman and Hulburt.

Table 1. Values of n'/n_0' for various γ's

h	n'/n_0'			
(km.)	$\gamma=0{\cdot}00$ (deg/km.)	$\gamma=2{\cdot}02$ (deg/km.)	$\gamma=4{\cdot}05$ (deg/km.)	$\gamma=8{\cdot}10$ (deg/km.)
200	$1{\cdot}00$	$1{\cdot}00$	$1{\cdot}00$	$1{\cdot}00$
225	$3{\cdot}64\times10^{-1}$	$3{\cdot}44\times10^{-1}$	$3{\cdot}25\times10^{-1}$	$2{\cdot}92\times10^{-1}$
250	$1{\cdot}33\times10^{-1}$	$1{\cdot}32\times10^{-1}$	$1{\cdot}30\times10^{-1}$	$1{\cdot}22\times10^{-1}$
275	$4{\cdot}78\times10^{-2}$	$5{\cdot}54\times10^{-2}$	$5{\cdot}98\times10^{-2}$	$6{\cdot}25\times10^{-2}$
300	$1{\cdot}74\times10^{-2}$	$2{\cdot}53\times10^{-2}$	$3{\cdot}05\times10^{-2}$	$3{\cdot}64\times10^{-2}$
325	$6{\cdot}28\times10^{-3}$	$1{\cdot}22\times10^{-2}$	$1{\cdot}69\times10^{-2}$	$2{\cdot}27\times10^{-2}$
350	$2{\cdot}29\times10^{-3}$	$6{\cdot}24\times10^{-3}$	$9{\cdot}93\times10^{-3}$	$1{\cdot}52\times10^{-2}$

The effect of a temperature gradient γ on the molecular density may be seen from Table 1, which has been worked out assuming the values $h_0=200$ km., $T_0=400°$ K., $\boldsymbol{g}=900$ cm.sec^{-2}, $m=15\times1{\cdot}66\times10^{-24}$ gm. and $\boldsymbol{k}=1{\cdot}38\times10^{-16}$ erg. deg^{-1}.

1 (iii). *Intensity of Incident Radiation*

The intensity I of the incident radiation (assumed monochromatic) will decrease as it passes through the atmosphere. Suppose that the intensity is increased by an amount dI in increasing the height from h to $h+dh$. If χ, the solar zenith distance, is small, so that we may consider the earth's surface to be flat,

$$dI=\beta n' I\sec\chi\,.\,dh, \qquad \text{.......(7)}$$

where β is the atomic absorption coefficient. Substitution from equation (2), and subsequent integration yields

$$I=I_\infty\exp\left\{n_0'T_0\beta\sec\chi\int_\infty^h\frac{1}{T}\exp\left(-\frac{m\boldsymbol{g}}{\boldsymbol{k}}\int_{h_0}^h\frac{dh}{T}\right)dh\right\}, \qquad \text{......(8)}$$

in which I_∞ is the intensity of the beam before it has reached the earth's atmosphere and I the intensity at height h.

1 (iv). *Rate of Ion Production*

Suppose q is the number of ion pairs produced per cm^3 per second at height h. Then if w is the energy absorbed in ionizing one molecule, and assuming that all the energy absorbed produces ionization,

$$q=\frac{\beta n' I}{w}. \qquad \text{......(9)}$$

Hence, on substituting from (8),

$$q=\frac{\beta n_0' I_\infty}{w}\cdot\frac{T_0}{T}\exp\left\{-C\int_{h_0}^{h}\frac{dh}{T}+n_0' T_0\beta\sec\chi\int_{\infty}^{h}\frac{1}{T}\exp\left(-C\int_{h_0}^{h}\frac{dh}{T}\right)dh\right\}. \qquad \ldots\ldots(10)$$

With a linear temperature gradient, according to equation (3), this becomes

$$\begin{aligned} q&=q_b\\ &=\frac{\beta n_0' I_\infty}{w}\left(1+\gamma\frac{h-h_0}{T_0}\right)^{-(1+C/\gamma)}\exp\left\{-\frac{n_0'\beta T_0}{C}\sec\chi\left(1+\gamma\frac{h-h_0}{T_0}\right)^{-C/\gamma}\right\}. \end{aligned} \qquad \ldots\ldots(11)$$

On proceeding to the limit $\gamma=0$, as before, this reduces to

$$\begin{aligned} q&=q_a\\ &=\frac{\beta n_0' I_\infty}{w}\exp\left\{-\frac{C}{T_0}(h-h_0)-\frac{n_0'\beta T_0}{C}\sec\chi\exp\left[-\frac{C}{T_0}(h-h_0)\right]\right\} \end{aligned} \qquad \ldots\ldots(12)$$

which, in the present notation, agrees with Hulburt and with Chapman.

1 (v). *Layer Formation*

If ions are produced according to equation (11) or (12), there will be an equal number n of positive ions and electrons per cm³. The rate of increase of ion density is then given by the well-known equation

$$\frac{dn}{dt}=q-\alpha n^2, \qquad \ldots\ldots(13)$$

in which α represents the effective recombination coefficient.

Case (a).—It is instructive to consider first the solution of equation (13) for the constant temperature case. Substituting from equation (12),

$$\frac{dn}{dt}+\alpha n^2=\frac{\beta n_0' I_\infty}{w}\exp\left\{-\frac{C}{T_0}(h-h_0)-\frac{n_0'\beta T_0}{C}\sec\chi\exp\left[-\frac{C}{T_0}(h-h_0)\right]\right\}. \qquad \ldots\ldots(14)$$

According to Chapman, this equation cannot be solved explicitly. Hulburt, however, has pointed out that during the middle part of the day dn/dt is small, and may be assumed to be zero to obtain an approximate equation. This assumption is confirmed by the experimental results to be described in § 2 of this paper. Equation (14) then reduces to

$$n_a^{\,2}=B\exp\left\{-\frac{C}{T_0}(h-h_0)-\frac{1}{F}\exp\left[-\frac{C}{T_0}(h-h_0)\right]\right\}, \qquad \ldots\ldots(15)$$

where we have written

$$B=\beta n_0' I_\infty/\alpha w \qquad \ldots\ldots(16)$$

and

$$F=C\cos\chi/n_0'\beta T_0. \qquad \ldots\ldots(17)$$

Case (*b*).—A linear temperature gradient, given by equation (3), will now be considered. As the form of the equation is more complex than in case (*a*), we proceed at once to make the approximation $dn/dt = 0$, obtaining

$$n_b{}^2 = B\left(1+\gamma\frac{h-h_0}{T_0}\right)^{-(1+C/\gamma)} \exp\left\{-\frac{1}{F}\left(1+\gamma\frac{h-h_0}{T_0}\right)^{-C/\gamma}\right\}. \qquad \ldots\ldots(18)$$

Equations (15) and (18) give the value of the electron density at any point in the layer.

1 (vi). *Maximum of Electron Density*

The maximum-with-height value N of the electron density, n, may be found by equating dn/dh to zero. When this is done with equations (15) and (18) in turn, we find that the required conditions are respectively

$$\frac{1}{F}\exp\left\{-\frac{C}{T_0}(h-h_0)\right\} = 1 \qquad \ldots\ldots(19)$$

and

$$\frac{1}{F}\left(1+\gamma\frac{h-h_0}{T_0}\right)^{-C/\gamma} = 1+\frac{\gamma}{C}. \qquad \ldots\ldots(20)$$

When substituted back in the original equations, we have

$$N_a{}^2 = BF\exp(-1) \qquad \ldots\ldots(21)$$

and

$$N_b{}^2 = B\left\{F\left(1+\frac{\gamma}{C}\right)\right\}^{(1+\gamma/C)} \exp\left\{-\left(1+\frac{\gamma}{C}\right)\right\}, \qquad \ldots\ldots(22)$$

giving the required maxima for the two cases.

The height of maximum electron density, h_M, is given by equations (19) and (20), which, for future reference, are now written as

$$(h_M)_a = h_0 - \frac{T_0}{C}\ln F, \qquad \ldots\ldots(23)$$

$$(h_M)_b = h_0 + \frac{T_0}{\gamma}\left[\left\{F\left(1+\frac{\gamma}{C}\right)\right\}^{-\gamma/C} - 1\right]. \qquad \ldots\ldots(24)$$

1 (vii). *Comparison of the Two Cases*

The changes in behaviour brought about by the assumption of a linear temperature gradient with height may be illustrated by means of the ratios

$$\frac{N_b}{N_a} \quad \text{and} \quad \frac{(h_M)_b - h_0}{(h_M)_a - h_0}.$$

By this method the use of a number of questionable constants may be eliminated. The theory assumes that the composition of the atmosphere at the reference height h_0 should be the same as that throughout the layer. There is reason to believe that the atmospheric gases exist mainly in atomic form from 200 km. upwards (Hulburt 1939 b); h_0 has therefore been chosen as 200 km., because the value of T_0 becomes less certain as the height increases. Further, putting $m = 15$ atomic weight units, $\boldsymbol{g} = 900$ cm.sec^{-2} and $\boldsymbol{k} = 1{\cdot}38\times10^{-16}$ erg.deg^{-1}, the value of C becomes 16·2 deg.km^{-1}. Finally, for purposes of this comparison only, assuming $T_0 = 360^\circ$ K., $n_0' = 10^{11}$ cm^{-3} and $\beta = 10^{-17}$ cm^2, on very debatable grounds, and putting $\chi = 0$ for simplicity, we find $F = 4{\cdot}4\times10^{-2}$.

Dividing equation (22) by (21), we obtain

$$\left(\frac{N_{\mathrm{b}}}{N_{\mathrm{a}}}\right)^2 = F^{\gamma/C}\,.\left(1+\frac{\gamma}{C}\right)^{(1+\gamma/C)}.\exp(-\gamma/C), \qquad \ldots\ldots(25)$$

and from equations (23) and (24)

$$\frac{(h_{\mathrm{M}})_{\mathrm{b}}-h_0}{(h_{\mathrm{M}})_{\mathrm{a}}-h_0} = -\frac{C}{\gamma}\,\frac{F\left(1+\frac{\gamma}{C}\right)^{-\gamma/C}-1}{\ln F}. \qquad \ldots\ldots(26)$$

To obtain some idea of the absolute difference in height of the maxima, we use

$$(h_{\mathrm{M}})_{\mathrm{b}}-(h_{\mathrm{M}})_{\mathrm{a}} = T_0\left[\frac{1}{\gamma}\left\{\left[F\left(1+\frac{\gamma}{C}\right)\right]^{-\gamma/C}-1\right\}+\frac{1}{C}\ln F\right], \qquad \ldots\ldots(27)$$

obtained by subtracting equation (23) from equation (24).

The values obtained are shown graphically in Figures 1 and 2. These graphs show that a positive value of γ produces a lowering of the maximum electron density in the layer and raises its height.

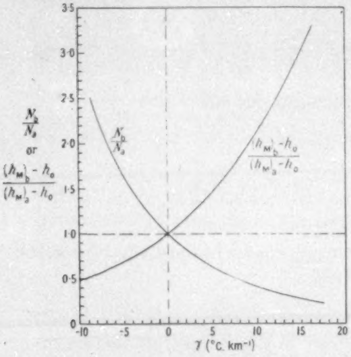


Figure 1.

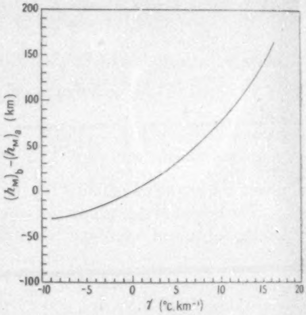


Figure 2.

1 (viii). *Effect of Variables on the Shape of the Layer*

As the general theory, summarized in equations (18), (22) and (24), does not lend itself to easy physical interpretation, it has been thought advisable to represent graphically the variations in shape of the layers brought about by changes in the variables. In order to eliminate as far as possible numerical errors arising from the use of uncertain values of the constants, equations (18) and (22) may be divided, giving

$$\left(\frac{n_{\mathrm{b}}}{N_{\mathrm{b}}}\right)^2 = \left(1+\gamma\frac{h_{\mathrm{b}}-h_0}{T_0}\right)^{-(1+C/\gamma)}\left\{F\left(1+\frac{\gamma}{C}\right)\right\}^{-(1+\gamma/C)}$$

$$\times\exp\left\{\left(1+\frac{\gamma}{C}\right)-\frac{1}{F}\left(1+\gamma\frac{h_{\mathrm{b}}-h_0}{T_0}\right)^{-C/\gamma}\right\} \qquad \ldots\ldots(28)$$

in which B does not appear. On the right-hand side of equation (28) we are thus left with four parameters, the height h_b, the temperature T_0 at 200 km., the temperature gradient γ, and F. In Figures 3, 4 and 5, $n_{\mathrm{b}}/N_{\mathrm{b}}$ is plotted against h_{b} for various values of γ, T_0 and F, using equation (28).

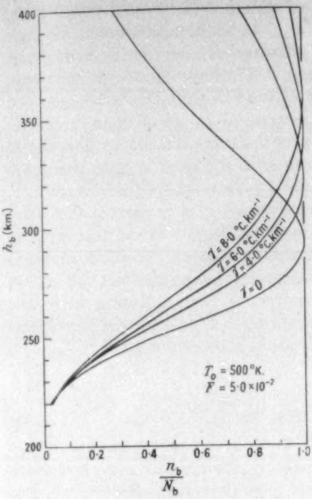


Figure 3. Effect of γ on shape of layer.

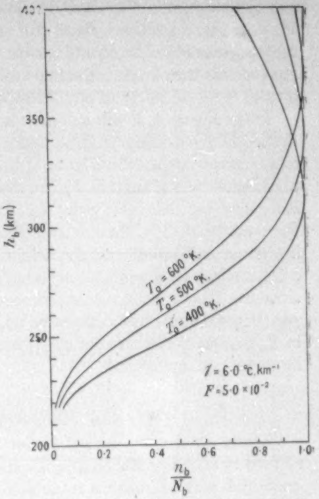


Figure 4. Effect of T_0 on shape of layer.

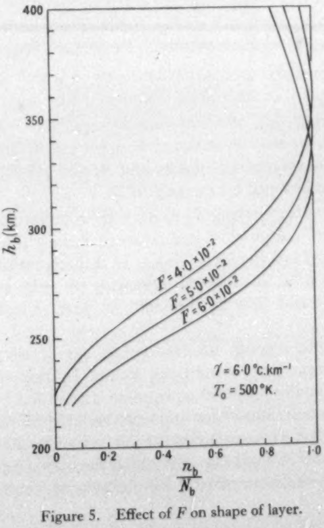


Figure 5. Effect of F on shape of layer.

In Figure 3, the curve for $\gamma=0$ is drawn according to Chapman's theory, and represents the fractional distribution of electrons at constant temperature. We note that a positive γ raises the height h_M. In addition, it is seen to increase the thickness of the layer, and the electron density above the maximum decreases more slowly than in an isothermal atmosphere. The importance of a knowledge of true layer thickness in any attempt to estimate γ is thus clearly brought out.

From Figure 4, it will be seen that an increase in T_0 produces an increase in the height of maximum electron density, but the thickness of the layer is not altered to any great extent. In Figure 5, the thickness of the layer is again practically unchanged by a change in F, but the height h_M decreases with increase in F. Remembering that for a given T_0 (and therefore n_0') F is proportional to $\cos\chi$ from equation (17), these curves show a result analogous to the theory for an isothermal atmosphere, i.e. the height of maximum electron density decreases to a minimum at local noon, all the other parameters remaining constant. However, since the theory of expansion demands atmospheric heating during the middle part of the day, T_0 and γ cannot be taken as independent of time. If the increase in T_0 is greater than that of $\cos\chi$, it alone would account for a rise in h_M, even if γ remained unchanged.

§ 2. EXPERIMENTAL DATA

Before developing a method by which the theory just expounded may be applied to experimental results, it is convenient to consider how many and which quantities may be derived from ionospheric measurements. Records of ionospheric conditions are almost universally given in the form of graphs of virtual height plotted against frequency; these are obtained either by automatic equipment such as that developed by Berkner, Wells and Seaton (1936) at the Carnegie Institute, or by a more modest form of manually operated equipment, where financial resources are not readily available or the proposed experimental work is not of a very permament nature. From such records may be obtained:

(i) the maximum electron density N in the layer, which is directly proportional to the square of the critical frequency for the ordinary ray;

(ii) the true height, h_M, of maximum density, and semi-thickness τ of the layer, which may be obtained without much difficulty either by the method developed by Booker and Seaton (1940) or by a more recent method developed by Pierce (1947).

At this stage it is interesting to remark how investigators have, up to the present, insisted on jumping to conclusions, basing their reasoning entirely on critical frequencies and virtual heights. In dealing with a problem as complex as the ionosphere, it is an obvious advantage to rely as far as possible on experimental results and thus to eliminate, in part at least, some of the little known parameters.

The photographic records used for purposes of testing this theory and gaining some information of conditions in the F_2 region were taken in this laboratory with manually operated equipment assembled by the writers during the war. A pulse transmitter of continuously variable frequency covered a range from 1·8 to 13·9 Mc/s.; its mean power output was about 200 watts during pulses. Records were made on 35 mm. film and scaled through a microfilm reader giving an approximate magnification of 10. By the help of frequency marks made on the record at regular intervals of 0·2 or 0·5 Mc/s. by means of a height calibrating oscillator working at 1,500 c/s., frequencies could be read to 0·01 Mc/s. and heights could be reproduced to 2 km.

A series of photographic records was taken, one on each hour on alternate days between 1st November 1945 and 28th February 1946, the summer season in South Africa; this choice of intervals was dictated by the shortage of observing staff available. All the records were completely scaled according to the method of Booker and Seaton (1940); heights were read at the 8 frequencies suggested by Booker and Seaton, and coupled in four pairs, each of them giving the true

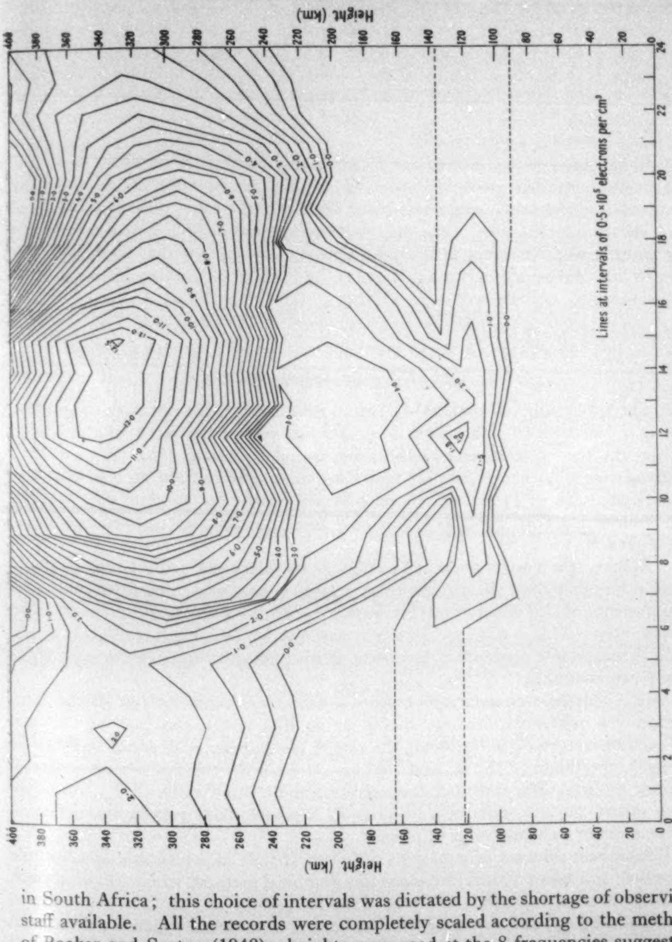


Figure 6. Electron density map. February 1946.

height of maximum density, h_M, and semi-thickness, τ, of the layer. The mean of the four values was calculated for each quantity. Where necessary, corrections due to retardation by lower layers were applied (Booker and Seaton 1940). Since a significant correlation was found between magnetically disturbed and ionospherically disturbed days, all days with magnetic character figure 2 were eliminated in taking the monthly mean values of N, h_M and τ. As some large deviations still remained, the more outstanding ones were removed, using the well established test that deviations larger than three times the probable deviation are unlikely to be due to chance alone. With the adjusted monthly mean values for N, h_M and τ thus obtained, it was possible to draw the Booker and Seaton parabolic distribution of (electron density, true height) for each hour of the average day for a given month.

It has been found convenient to represent all the information about the behaviour of the ionosphere on the average day for each month by a graph showing lines of constant n on a coordinate plane of time and height. This plot is called an 'electron density map': the map for the February data is shown in Figure 6 as an example. It serves to illustrate the important point that dn/dt is in fact nearly zero during the daytime in parts of the F_2 layer below the electron density maximum.

§ 3. APPLICATION OF THE THEORY TO EXPERIMENTAL RESULTS

3 (i). *The Method of applying the Theory*

The application to experimental results of a theory with equations as complex as those derived in § 1 of this paper is a problem which must present some difficulty. From the theory there emerge the three useful equations (18), (22) and (24). Considering C as known, we see that these equations contain no less than four unknowns: B, γ, T_0 and F. B may be eliminated by dividing (18) by (22), thus obtaining (28), but this leaves only two equations with three unknowns, γ, T_0 and F.

At first sight it might seem that by taking two values of h and the corresponding two values of n from the experimental parabola, we could obtain two numerical equivalents of (28) which, together with (24), would allow solution for all three unknowns. This is a fallacy, however, because by dividing equations (18) and (22), the absolute value of N has been eliminated and the parabola now yields only two quantities.

It is therefore necessary to assume a numerical value for one of the three remaining unknowns, the one chosen being that which can best be assessed. The temperature T_0 is obviously the one to be chosen, as it is not likely to be outside the limits of 350° K. and 700° K., i.e. it varies through no more than a factor of two. We shall not, however, assume a fixed value of T_0, but apply the theory for a number of values of T_0 between these two limits, deferring discussion of its actual value until later.

Algebraic solution of equations (24) and (28) is impracticable as a routine method, and resort is therefore made to a graphical method, which can be applied quickly to most conditions met with in the F_2 region. It should be remembered that the final equations are only true if $dn/dt = 0$, a condition which is well fulfilled for the F_2 region below its maximum from 08h. to 16h. South African Standard Time, as previously pointed out.

3 (ii). *Graphical Solution for F and γ*

The value of n for substitution in equation (28) may be read at any height below h_M, where the electron density is greater than $0{\cdot}3N$. These limits are those beyond which the curve differs appreciably from parabolic form, and between which all the virtual heights are originally scaled.

In order to construct a family of curves of $(n/N, \gamma)$ for various values of F from equation (28), the value of $(h-h_0)/T_0$ must be settled. A value of h agreeing with the conditions outlined in the previous paragraph is 280 km., as may be verified by an examination of the electron density map. Assuming that T_0 will be of the order of 400° K., we find

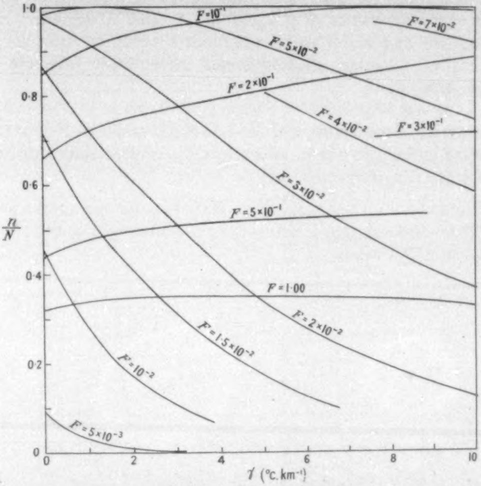


Figure 7. Curves of $\frac{n}{N}$ against γ for $\frac{h-h_0}{T_0} = \frac{1}{5}$, from equation (30).

$$\frac{h-h_0}{T_0} = \frac{280-200}{400} = \frac{1}{5}. \qquad \ldots\ldots(29)$$

When we recall that $C = 16{\cdot}2$° c. km^{-1}, and substitute in (28), dropping the subscript b, this simplifies to

$$\left(\frac{n}{N}\right)^2 = \left(1+\frac{\gamma}{5}\right)^{-(1+16\cdot2/\gamma)} \left\{F\left(1+\frac{\gamma}{16\cdot2}\right)^{-(1+\gamma/16\cdot2)} \times \exp\left\{\left(1+\frac{\gamma}{16\cdot2}\right) - \frac{1}{F}\left(1+\frac{\gamma}{5}\right)^{-16\cdot2/\gamma}\right\}\right. . \qquad \ldots\ldots(30)$$

Choosing values of γ which give simple numerical values to the indices ($\gamma = 1{\cdot}62$, $3{\cdot}24$, $4{\cdot}86$, $6{\cdot}49$, $8{\cdot}10$, $9{\cdot}7$), and using equations (15) and (21) for $\gamma = 0$, various values of F may be substituted to give n/N greater than $0{\cdot}3$. In this way the family of curves shown in Figure 7 was drawn.

Equation (24) may be written, with the usual value of C,

$$\frac{h_M - h_0}{T_0} = \frac{1}{\gamma}\left[\left\{F\left(1+\frac{\gamma}{16{\cdot}2}\right)\right\}^{-\gamma/16\cdot 2} - 1\right]. \qquad \ldots\ldots(31)$$

With the values of γ used before a similar family of curves may be drawn, and is shown in Figure 8.

These curves exhibit a number of interesting points:

(i) The occurrence of maxima in the curves of n/N for $F = 5 \times 10^{-2}$ and $F = 10^{-1}$ for values of γ equal to 1·2 and 8 deg.km^{-1} respectively indicate that under these conditions n becomes equal to N, i.e. the maximum electron density passes through the height h for which $(h-200)/T_0 = \frac{1}{5}$.

As was illustrated by Figure 5, an increase in F corresponds to a lowering of the height, and the curves for values of F greater than 10^{-1} refer entirely to values of n above h_M, and therefore cannot be used for practical purposes.

(ii) In Figure 8 the curves below $(h_M - 200)/T_0 = 0{\cdot}2$ again refer to conditions when the maximum is below the selected height h, and therefore have no practical value.

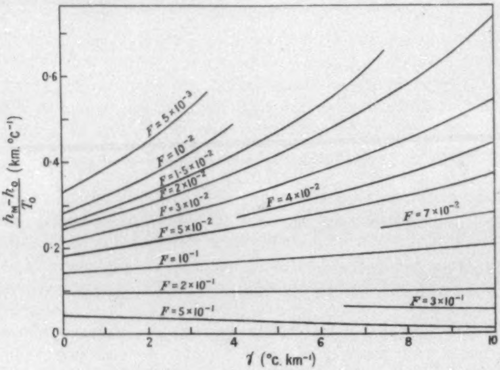


Figure 8. Curves of $\frac{h_M - h_0}{T_0}$ against γ for $\frac{h-h_0}{T_0} = \frac{1}{5}$, from equation (31).

The simultaneous graphical solution of equations (30) and (31) may now be accomplished, if values taken both from Figure 7 and from Figure 8 are plotted on the same diagram, using as coordinates the variables F and γ or, preferably, $\log F$ and γ. This has been done in Figure 9, where lines of constant n/N from Figure 7 are drawn in full and those of constant $(h_M - 200)/T_0$ from Figure 8 are shown broken.

The process of evaluating γ and F for a given experimental parabola now resolves itself into the following series of simple operations:

(i) select a reasonable value of T_0;
(ii) estimate the value of h by substituting this value of T_0 in the equation $(h-200)/T_0=\frac{1}{5}$;
(iii) from the parabola, find n at this value of h, and also N;
(iv) hence calculate n/N;
(v) from the parabola read off h_M;
(vi) evaluate $(h_M-200)/T_0$, using the value of T_0 chosen in (i);
(vii) on Figure 9 locate the point which satisfies simultaneously (iv) and (vi), and read off the values of γ and $\log F$.

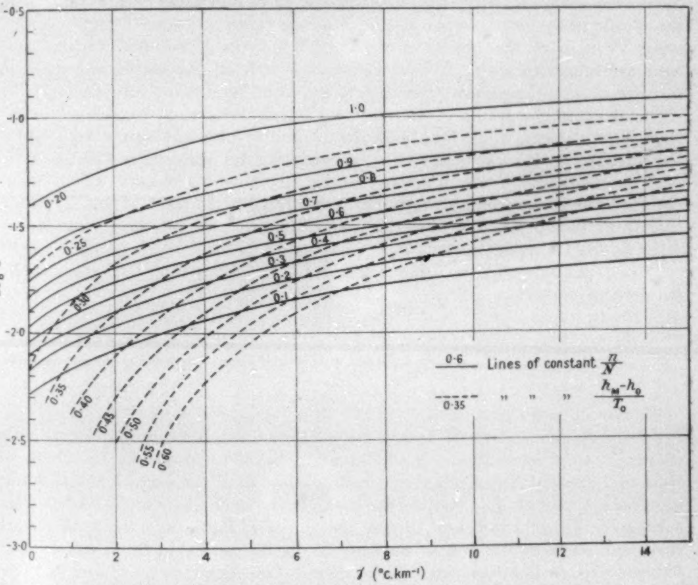


Figure 9. Graph for the evaluation of γ and F from experimental data.

The whole procedure is repeated for each value of T_0 selected.

The curve $n/N=1{\cdot}0$ represents the maximum electron density, and thus the part of the graph above this is not shown, as it has no use. It coincides with the line $(h_M-200)/T_0=0{\cdot}2$, which also represents the electron density maximum.

The angle between the two families of curves increases as we estimate n further below the maximum, and therefore values of γ and F estimated from values of n scaled near the maximum are not very accurate. This could be avoided by choosing the fraction $(h-200)/T_0 \neq \frac{1}{5}$, but this would introduce other difficulties by putting n below $0{\cdot}3N$ in the normal applications.

The two families of curves tend to run parallel for values of γ greater than 15 deg.km^{-1}, and this introduces uncertainty in values estimated in this region. Since for $\gamma = 15$ deg.km^{-1} and $T_0 = 250°$ K. (a very low estimate), the temperature at 350 km. would be 2,500° K., which is more than anyone has hitherto dared to suggest, this is no disadvantage. Godfrey and Price (1937) have shown by considerations of radiation equilibrium that this temperature cannot rise above 3,300° K.

3 (iii). *Treatment of Experimental Results*

The process just developed was applied to the experimental results for November 1945–February 1946, described in § 2 of this paper. The value of T_0 being unknown, the estimation of γ and F was repeated for each parabola for a series of values of T_0, varying in steps of 50° K. from 250 to 600° K. In many cases the higher values of T_0 gave values of γ well off the graph, and these, in accordance with the previous discussion, have not been taken into account.

To illustrate the results of this treatment and the following discussion, Table 2 is given. It shows, in the first two rows of each section, the figures for γ and F for November. Similar tables were prepared for the other three months.

The values of γ obtained are practically all above 5 deg.km^{-1} which points straight away to temperatures of the order of 1,000° K. in the F_2 layer. Further discussion on this point is, however, premature, because we have no means of selecting any particular value of T_0 as yet.

Some indication of the correct value of T_0 and γ may be obtained by performing two tests.

Test I.

From equation (17) we find

$$\beta = \frac{C\cos\chi}{n_0' T_0 F}. \qquad \ldots\ldots(32)$$

As the maintenance of the layer is presumably due to ionization of a particular species of atom or molecule, β should be constant. Cos χ may be calculated from the Nautical Almanac and we have a value of F for each T_0. However, n_0' will also vary with T_0. During the preliminary work this was calculated from a probable value at 100 km. on the assumptions that there the temperature remained unchanged at 360° K. throughout the day and that there was a linear temperature gradient between the 100 and 200 km. levels. Later a more detailed examination of the problem gave what is believed to be a better value for the 200 km. level, $n_0' = 1{\cdot}5 \times 10^{11}$ cm^{-3} when $T_0 = 360°$ K., but did not indicate the variation of this quantity with T_0. For the present purposes this quantity has been taken as a datum and modified for other values of T_0 by assuming it to vary in the same proportion as the earlier estimate. The values so obtained are shown in Table 3.

In this way the values of β corresponding to each T_0 at each hour were evaluated, and those for November are given in the third row of each section of Table 2.

Table 2. Experimental values of γ and F and derived values of β and B/n'_0. November 1945

S.A.S. Time	T_0 (° K.)	250	300	350	400	450	500	550	600
	γ	11·7	10·7	9·2	8·5	8·0	9·5	11	—
8 h.	100F	3·94	4·34	4·68	5·25	5·85	7·50	9·33	—
$\cos\chi=0{\cdot}529$	$\beta\times10^{17}$	10·6	6·04	3·63	2·41	1·82	1·17	0·78	—
	B/n'_0	4940	2740	1440	901	675	529	396	—
	γ	11·4	9·9	8·9	7·9	7·9	6·5	10	—
9 h.	100F	3·71	4·02	4·49	4·87	5·70	5·89	8·71	—
$\cos\chi=0{\cdot}701$	$\beta\times10^{17}$	15·0	8·65	5·00	3·45	2·48	1·97	1·10	—
	B/n'_0	6030	3120	1660	1080	815	565	461	—
	γ	13·5	12·7	11·6	10·7	9·2	11·5	13·0	—
10 h.	100F	4·36	5·01	5·46	5·92	6·03	7·76	9·12	—
$\cos\chi=0{\cdot}836$	$\beta\times10^{17}$	18·1	8·27	4·88	3·39	2·80	1·78	1·27	—
	B/n'_0	8170	4240	2370	1550	1160	942	773	—
	γ	11·3	9·65	8·3	7·6	7·1	7·0	5·5	—
11 h.	100F	3·21	3·36	3·61	4·01	4·52	5·21	5·25	—
$\cos\chi=0{\cdot}927$	$\beta\times10^{17}$	22·8	13·7	8·25	5·56	4·12	2·94	7·73	—
	B/n'_0	13100	6790	3550	2290	1690	1220	890	—
	γ	10·2	9·3	8·0	7·2	6·6	6·7	5·5	9·8
12 h.	100F	2·69	3·14	3·36	3·76	4·20	4·98	5·07	8·32
$\cos\chi=0{\cdot}966$	$\beta\times10^{17}$	28·3	15·3	9·20	6·18	4·65	3·21	2·62	1·37
	B/n'_0	17200	8470	4520	2860	2110	1510	1110	920
	γ	8·5	8·2	7·1	6·0	6·0	5·5	4·0	9
13 h.	100F	2·11	2·67	2·92	3·19	3·82	4·22	4·17	7·94
$\cos\chi=0{\cdot}952$	$\beta\times10^{17}$	35·8	17·6	10·4	7·17	5·01	3·74	3·12	1·40
	B/n'_0	15600	8070	4210	2600	1920	1390	993	812
	γ	9·8	8·9	7·5	6·8	6·1	5·4	3·6	8·6
14 h.	100F	2·69	3·10	3·29	3·73	4·13	4·44	4·09	8·03
$\cos\chi=0{\cdot}884$	$\beta\times10^{17}$	25·6	14·1	8·56	5·70	4·31	3·29	2·96	1·29
	B/n'_0	13300	6990	3680	2320	1690	1210	895	720
	γ	7·8	7·4	6·3	5·6	4·8	3·8	8	—
15 h.	100F	2·25	2·79	3·09	3·55	3·86	4·05	7·59	—
$\cos\chi=0{\cdot}769$	$\beta\times10^{17}$	27·0	13·6	7·96	5·19	4·01	3·13	1·39	—
	B/n'_0	11200	5770	3000	1900	1370	960	726	—
	γ	6·5	6·1	5·2	4·8	3·6	5	12	—
16 h.	100F	2·01	2·57	2·98	3·52	3·69	5·62	10	—
$\cos\chi=0{\cdot}613$	$\beta\times10^{17}$	24·0	11·8	6·59	4·07	3·35	1·81	0·84	—
	B/n'_0	9020	4520	2310	1440	1040	718	689	—

Table 3. Values of n'_0 for various T_0

T_0 (° K.)	250	300	350	400	450	500	550	600
$n'_0\times10^{-11}$ cm^{-3}	0·935	1·20	1·45	1·69	1·90	2·07	2·27	2·44

Test II.

The second test depends on the evaluation of the constant B, previously eliminated by division of equations (18) and (22). B is defined by equation (16) and is also given by a rearrangement of equation (22). From its definition it is seen that B is not a constant, since it contains n_0' which is a function of T_0. The quantity

$$\frac{B}{n_0'} = \frac{\beta I_\infty}{\alpha w} \quad \text{from equation (16)}$$

$$= \frac{N^2 \exp(1+\gamma/C)}{n_0' \left\{ F\left(1+\frac{\gamma}{C}\right) \right\}^{(1+\gamma/C)}} \quad \text{from equation (22)} \qquad \ldots\ldots(33)$$

should however be constant at all times and for all months. The values for November are shown in the fourth row of each section of Table 2.

3 (iv). *Discussion of the Tests*

The correct value of β is very doubtful. Table 2 and the corresponding tables for the other three months have been drawn up to cover temperatures T_0 from 250° K. to 600° K., which range includes most of the rough estimates found in the literature, and, if this range is accepted as the most probable for T_0, the correct value of β should appear in the third row of each section. An examination of the tables shows that values of β in the neighbourhood of 3×10^{-17} cm^2 fulfil this condition. Table 4 shows the values of T_0, interpolated to the nearest 10° K. from the tables, which correspond to $\beta = 3{\cdot}2 \times 10^{-17}$ cm^2;

Table 4. Values of T_0 (° K.) corresponding to $\beta = 3{\cdot}2 \times 10^{-17}$

S.A.S. Time (hours)	November	December	January	February
8	370	320	320	320
9	410	380	350	360
10	410	410	470	390
11	480	540	410	430
12	500	490	470	520
13	540	460	430	450
14	500	450	420	440
15	490	430	390	410
16	450	390	330	390

In theory, if a value of B/n_0' is chosen to give the same T_0 as the value $\beta = 3{\cdot}2 \times 10^{-17}$ cm^2 at a certain time, this same value of B/n_0' should correspond to the identical set of values of T_0 shown in Table 4. However, B contains α and I_∞, which may vary with season, sunspot activity, etc., respectively. Hence we shall not expect exact agreement in practice. To find the value of B/n_0' which best corresponds to $\beta = 3{\cdot}2 \times 10^{-17}$ cm^2, a figure was estimated to correspond to each of the noon values of T_0 in Table 4, and the mean of the four was taken. This gave the best value of B/n_0' as 910 cm^{-3}; the agreement was not good, the largest deviation from the mean being about 50%. Considering the relatively large percentage change in B/n_0' corresponding to a change of only 50° K. in T_0, however, this may be regarded as satisfactory. The T_0 values corresponding to $B/n_0' = 910$ cm^{-3}, interpolated for the last row of Table 2 and the corresponding ones for the other three months, are shown in Table 5.

Table 5. Values of T_0 (° K.) corresponding to $B/n_0' = 910$ cm^{-3}

S.A.S. Time (hours)	November	December	January	February
8	400	330	300	320
9	430	330	330	400
10	510	350	350	450
11	550	440	350	450
12	600	410	370	550
13	570	410	350	560
14	550	420	350	550
15	510	390	340	550
16	470	380	310	460

The agreement between the two tests is very pleasing, as they rest on independent grounds; therefore it does not seem out of place to take the mean values from these two tables the best representative values for further discussion. These are given in Table 6.

Table 6. Mean values of T_0 (° K.)

S.A.S. Time (hours)	November	December	January	February
8	385	325	310	320
9	420	355	340	380
10	460	380	410	420
11	515	490	380	465
12	550	450	420	535
13	555	435	390	505
14	525	435	385	495
15	500	410	365	480
16	460	380	320	425

The absolute values of these T_0's depend, of course, on the arbitrary choice of β as $3{\cdot}2 \times 10^{-17}$ cm^2. The effect of changing this was investigated by estimating new values of T_0 by the two tests with $\beta = 1{\cdot}6 \times 10^{-17}$ cm^2 and $8{\cdot}0 \times 10^{-17}$ cm^2 respectively, and the corresponding figures for B/n_0'. This resulted in an increase of the estimated T_0 by about 100° K. in the first case and a decrease by the same amount in the second case, but it did not affect the form of the variation of T_0 with time of day, as may be easily verified by an examination of Table 2.

Previous estimates have been made of the mean temperature at 200 km., but they vary widely in order of magnitude. Martyn and Pulley (1936), from considerations of electron collision frequencies and of auroral spectra, have tentatively adopted the value of 1,000° K. at 200 km. in the daytime; Hulburt (1939 b) has assumed the value 360° K. above 100 km. for lack of better evidence; while Vassy and Vassy (1942), assuming complete dissociation of both oxygen and nitrogen, obtained a temperature gradient which gives T_0 at 200 km. as 600° K. in summer. The values in Table 6, having a mean of about 430° K., are well in accord with these estimates.

It may be seen that T_0 rises towards the middle of the day and falls off with some delay in the afternoon. This is true no matter what β is chosen to be.

It shows that the heating of the atmosphere does take place even at a height of 200 km., the temperature rising by about 30% of its value at 8 hours.

Estimates have been made of the 'temperature of the F_2 region' during summer daytime:

(i) Appleton and Naismith (1935), from considerations of the variation in N between summer and winter noon, estimate that the temperature cannot be less than 1,200° K. in summer.

(ii) Fuchs (1936) has concluded from estimated scale heights that the temperature of the F_2 region may rise as high as 1,900° K.

(iii) Martyn and Pulley (1936) confirm Appleton's result of a temperature of 1,200° K.

(iv) Vassy and Vassy (1942) assuming complete dissociation of O_2 and N_2, obtain a summer temperature of 1,080° K. from the estimated scale height of 70 km. given by Appleton.

(v) Godfrey and Price (1937), by a consideration of the equilibrium between absorption and secondary radiation, conclude that the temperature is of the order of 1,000° K., and cannot exceed 3,300° K.

For comparison, the temperature at 300 km. was calculated from the T_0's of Table 6 and the corresponding γ's. The individual variations of these figures for T_{300} are probably not significant of actual conditions, and therefore only the means are included in Table 7.

Table 7. Mean Temperatures at 200 and 300 km.

Month	November	December	January	February
Mean T_0 (° K.)	485	408	369	447
Mean T_{300} (° K.)	1097	1376	1392	1305

It will be observed that the mean values for T_0 are smaller for the midsummer months December and January than for November and February, while the reverse holds for temperatures at 300 km. This gives valuable support to the expansion theory, for, near the solstice, when heating is greatest, the atmosphere at 200 km. becomes heated in the middle part of the day, and is forced upwards by expansion, leaving cooler gas to take its place. The heated mass of gas would produce greater absorption of the solar radiation, thus preventing it from reaching the 200 km. level, where it would normally cause more heating.

As a last interesting point, we may make an estimate of the intensity of active incident radiation I_∞ above the earth's atmosphere. From equation (16)

$$I_\infty = \left(\frac{B}{n_0'}\right)\frac{\alpha w}{\beta}. \qquad \text{......(34)}$$

We have assigned to B/n_0' the value 910 cm^{-3} from the results of test I. The value of α is open to discussion, and seems to vary with ion density and height, according to Wells and Shapley (1946); a reasonable mean seems to be 5×10^{-10} $\text{cm}^3\,\text{sec}^{-1}$; β has been fixed at $3{\cdot}2 \times 10^{-17}$ cm^2; and the value of the ionization potential w is given by Bacher and Goudsmit (1932) as 13·55 ev. for atomic oxygen and 14·48 ev. for atomic nitrogen. As a mean 14 ev. has been used. Thus we find

$$I_\infty = 0{\cdot}313 \text{ erg.cm}^{-2}\,\text{sec}^{-1}.$$

Hulburt (1939 b) has calculated that the total black body radiation of energy per quantum greater than 13·7 ev. received from the sun is 8×10^{-3} erg.cm^{-2}sec^{-1}, which is much lower than the above estimate.

Substituting back the value 8×10^{-3} in equation (34), we find $B/n_0' = 23{\cdot}2$, and from Table 2 this will be seen to correspond to T_0 at least as high as 2,000° K. This is ridiculous, and we can only conclude that the sun does not behave as a black body in this spectral region. A similar conclusion has been drawn by other workers (Bradbury 1938, Saha 1937, McNish 1937, Hunter 1943).

§ 4. CONCLUSIONS

The introduction of a linear temperature gradient into the theory of layer formation yields equations which may be applied with success for data for the F_2 layer, presented in the form of parabolae representing the distribution of ion density with height. The values of the temperatures at 200 and 300 km. obtained are in accordance with previous estimates, and are of a more quantitative nature.

The main weakness seems to lie in the assumption regarding the absorption coefficient β. In the first place its value is not fixed by the theory, and in the second place it varies with the frequency of the radiation absorbed, and an assumption of a mean value is not necessarily justified. A general treatment of this problem has been given by Chapman (1939). Again, the manner in which the recombination coefficient α varies with the temperature and pressure is not well established. Finally the theory is limited in its application by the assumption that $dn/dt = 0$ and by its neglect of the earth's curvature. Re-examination of these points seems to be the next step in its development.

It should be pointed out that the assumption of complete dissociation of oxygen and nitrogen above 100 km. is not a weakness, since a numerical recalculation is all that is required to apply the theory to an atmosphere containing constituents of known molecular weight.

ACKNOWLEDGMENTS

In conclusion it is a pleasure to thank Professor Varder for his constant encouragement and for making available the resources of his department, and the National Research Board and Council for a grant.

REFERENCES

Appleton, E. V., and Naismith, R., 1935, *Proc. Roy. Soc.* A, **150,** 685.
Bacher, R. F., and Goudsmit, S., 1932, *Atomic Energy States* (New York : McGraw-Hill).
Berkner, L. V., Wells, H. W., and Seaton, S. L., 1936, *Trans. Edin. Meeting Assn. Terr. Magn. Elect.*, p. 340.
Booker, H. G., and Seaton, S. L., 1940, *Phys. Rev.*, **57,** 87.
Bradbury, N. E., 1938, *Terr. Magn. Atmos. Elect.*, **43,** 55.
Chapman, S., 1931, *Proc. Phys. Soc.*, **43,** 26, 483; 1939, *Ibid.*, **51,** 93.
Fuchs, J., 1936, *Beitr. Geophys.*, **47,** 1.
Godfrey, G. H., and Price, W. L., 1937, *Proc. Roy. Soc.* A, **163,** 228.
Hulburt, E. O., 1928, *Phys. Rev.*, **31,** 1018; 1938, *Ibid.*, **53,** 344; 1939 a, *Ibid.*, **55,** 639; 1939 b, *Physics of the Earth*, edited by J. A. Fleming (New York : McGraw-Hill), Chap. X, p. 492.
Hunter, A., 1943, *Rep. Prog. Phys.*, **9,** 5 (London : Physical Society).
McNish, A. G., 1937, *J. Appl. Phys.*, **8,** 718.
Martyn, D. F., and Pulley, O. O., 1936, *Proc. Roy. Soc.* A, **154,** 455.
Millington, G., 1932, *Proc. Phys. Soc.*, **44,** 580; 1935, *Ibid.*, **47,** 263.
Pierce, J. A., 1947, *Phys. Rev.*, **71,** 698.
Saha, M. N., 1937, *Proc. Roy. Soc.* A, **160,** 155.
Vassy, A., and Vassy, M. E., 1942, *J. Phys. Radium*, ser. 8, **3,** 8.
Wells, H. W., and Shapley, A. H., 1946, *Terr. Magn. Atmos. Elect.*, **51,** 401.

A New Method of Measurement of the Variation with Wavelength of the Refractive Index and Absolute Stress Optical Coefficients of Amorphous Solids

By J. E. H. BRAYBON
University College, Southampton

Communicated by S. Weintroub; MS. received 9th January 1950

ABSTRACT. White light is passed through a Michelson interferometer in one of whose arms a thin lamina of the substance in the shape of a tension member is supported by a loading frame. The emergent beam, examined by a spectrometer, shows a continuous spectrum crossed by interference fringes. For the unstressed lamina the position of the fringes gives the variation of refractive index with wavelength. For the stressed lamina the behaviour of the fringes gives the stress optical coefficients. No modification of the interferometer, and only three sets of observations on the fringes, are required. For the calculation of absolute values of the refractive index the values of two wavelengths must be known.

Full details of the method are given, and photographs of tests on Columbia Resin 39 are included. These show that CR39 becomes positively birefringent under tension.

§ 1. INTRODUCTION

The method, which uses an unmodified Michelson interferometer, has been specially developed in order that measurements over a wide range of wavelength of both the refractive index and stress optical coefficients can be made on the same specimen. This is of importance in the case of a plastic, where the optical properties may vary with age and from sample to sample, and thus, unless the same specimen is used throughout, the refractive index cannot be correlated with the stress birefringence. The material used in checking the method was a condensed allyl ester, Columbia Resin 39 (CR39). In one arm of the interferometer, a thin lamina of the material (1 to 2 mm. thick), in the form of a tension member of uniform cross section, is held by a frame which may be used to apply a homogeneous tensile stress to the specimen. White light is passed through the system and the emergent beam is examined by a spectrometer. The continuous spectrum is observed to be crossed by a system of interference bands. From this system for the unstressed lamina, the variation of refractive index with wavelength is obtained; from the behaviour of the system when the lamina is stressed the stress optical coefficients are deduced. The theory of the method is given in detail.

§ 2. DETERMINATION OF REFRACTIVE INDEX DISPERSION CURVE

(i) *Experimental Details*

The material, in the form of a plane parallel thin lamina, is inserted in one arm of a Michelson interferometer. Light from a 'Pointolite' source P (Figure 1) is focused on an adjustable circular diaphragm I, collimated by the lens L and passes into the interferometer M. Emerging from M, it is focused as a vertical line by the cylindrical lens C on to the slit of a Hilger constant deviation spectrometer H, or a spectrograph.

If great care is taken to align all the components strictly along the optical axis of the system, then, for a suitable setting of the tilt of the interferometer mirror m_2, the resultant spectrum is observed to be crossed by vertical dark interference fringes. The relation between fringe wavelength λ, refractive index μ, and thickness t, of the lamina may be expressed as

$$n\lambda = D - 2(\mu - 1)t, \qquad \ldots\ldots(1)$$

where D is the optical path difference between the two interfering beams in the absence of the specimen and n, the fringe order, is an integer.

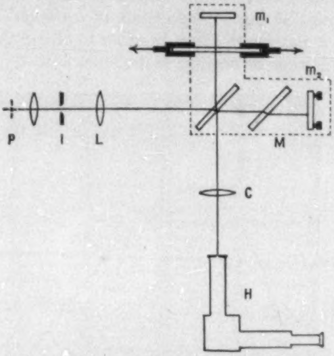


Figure 1. Optical system.

Due to the dispersion of μ with λ the fringes are not equidistant in wave number but are approximately symmetrical about a wavelength which is dependent upon D, i.e. upon the setting of the movable mirror, m_1, of the interferometer. The photograph, Figure 2 (Plate I), obtained with a lamina of CR39, thickness 1·68 mm., shows the variation in position of this central fringe with alteration of D by lateral movement of m_1. Further, replacing the lens C by a spherical lens, the diaphragm I by a vertical slit, and the source by a line filament, circular fringes are obtained and, in the region of the central fringe, it may be confirmed that corresponding fringes on either side of the central fringe are of the same order. Also, from equation (1), it is apparent that the fringe order rises to a maximum at the central fringe.

(ii) *Calculation of the* (μ, λ) *Dispersion Curve*

The absolute value of the fringe order cannot be determined directly without knowing the other variables in equation (1), but by plotting the fringe order (arbitrary) against λ, a curve of the form shown in Figure 3 can be obtained. From this curve the dispersion curve of μ against λ may be obtained as follows: Let the wavelengths of two fringes, on opposite sides of the maximum, but having the same fringe order, x (absolute value unknown), be λ_0 and λ_0'. If the wavelength of any fringe of order $(x+m)$ is λ_m, and μ_0, μ_0' and μ_m are the values of the refractive index at λ_0, λ_0' and λ_m respectively, then, from equation (1), by elimination of x and D:

$$\mu_0-\mu_m=(\mu_0-\mu_0')(\lambda_m-\lambda_0)/(\lambda_0'-\lambda_0)+m\lambda_m/2t. \qquad \ldots\ldots(2)$$

If μ_0 and μ_0' are known, μ_m is given by the above equation. However, it is not necessary in practice to calculate μ_m directly from equation (2), as the (μ,λ) dispersion curve can be obtained by a simple graphical construction.

In the absence of the $m\lambda_m/2t$ term, equation (2) represents the broken straight line of Figure 3 (*b*), through the points (λ_0, μ_0), (λ_0', μ_0'), which is easily constructed. By adding to the ordinate of any point on this line the corresponding value of $m\lambda_m/2t$ the corresponding point on the dispersion curve is at once obtained.

It should be noted that this interpretation of equation (2) confirms that the fringe order rises to a maximum, since, in order to obtain the normal dispersion curve, the term $m\lambda_m/2t$ must be positive in the range λ_0 to λ_0', and negative outside this range.

For this method, either μ_0 and μ_0' must be determined by other means, or D and one value of refractive index must be known. With the interferometer used

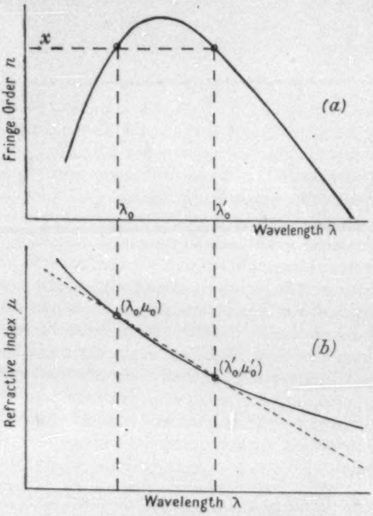


Figure 3. (*a*) Variation of fringe order (n) with wavelength λ.
(*b*) Variation of refractive index (μ) with wavelength λ.

it was more convenient to adopt the former course. Furthermore, unless the lamina itself is absolutely plane parallel, the interferometer, when the lamina is removed, will not give fringes without an alteration in the tilt setting of the mirrors in addition to a change in D. This completely precludes a determination of

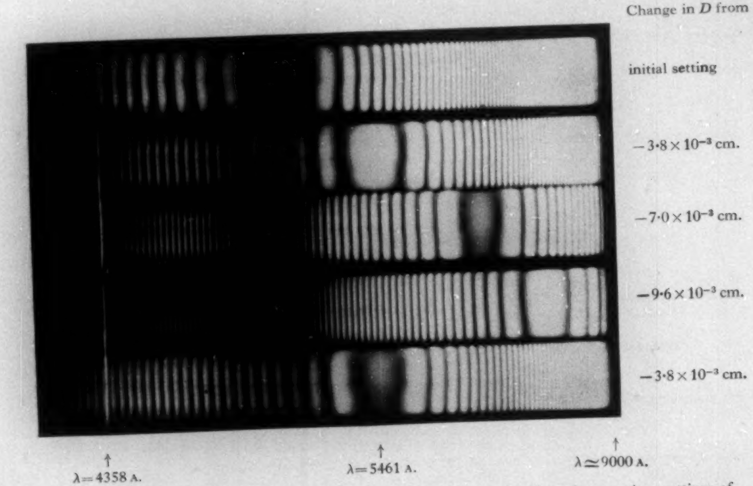

Change in D from
initial setting
$-3{\cdot}8\times10^{-3}$ cm.
$-7{\cdot}0\times10^{-3}$ cm.
$-9{\cdot}6\times10^{-3}$ cm.
$-3{\cdot}8\times10^{-3}$ cm.
↑ $\lambda=4358$ A.
↑ $\lambda=5461$ A.
↑ $\lambda\simeq9000$ A.

Figure 2. Photograph (positive print) of fringe systems obtained with various carriage settings of the Michelson interferometer. Tension member in position indicated in Figure 1.

PLATE I.

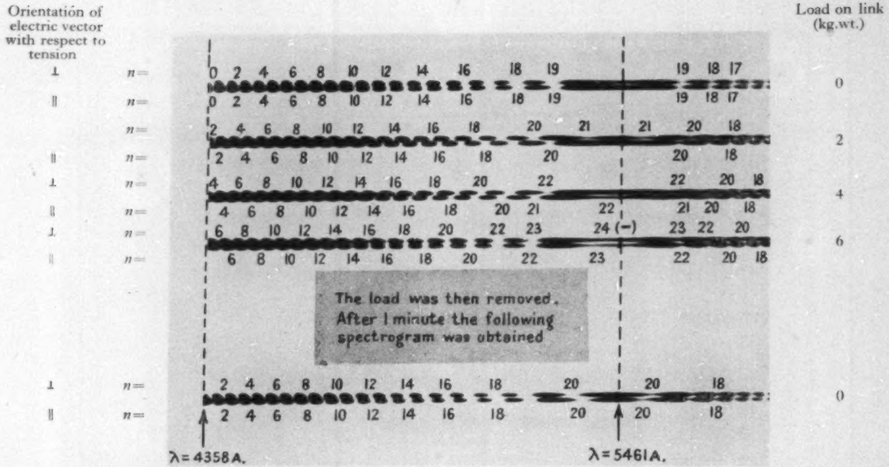


Figure 4. Enlarged (negative) portion of fringe systems obtained with various loads on link. Rochon prism in entry beam of interferometer.

PLATE II.

D since an adjustment of the mirror tilt in turn affects D. Because of this μ_0 and μ_0' were determined on two other specimens, both in the shape of 60° prisms. Experience had shown that in the case of CR39 the use of other samples is permissible as its optical properties are quite constant. Such a course, however, would not be suitable for a plastic such as Catalin, which is very susceptible to time effects, and the determination of μ must be made in this case directly by a refractometer or by a method involving the measurement of the Brewsterian angle (Pfund 1941). The final dispersion curve obtained agrees very closely with readings made using the 60° prisms at various other points over the visible range. The lamina used was supplied by the makers from their normal range, and, apart from cutting it into the form of a link for the purpose of stressing it as described below, it required no treatment.

§ 3. MEASUREMENT OF THE STRESS OPTICAL COEFFICIENTS

(i) *Experimental Details*

The lamina used must be in the form of a tension member which is subjected to a homogeneous tensile stress by means of a simple bell crank loading device. On loading as indicated in Figure 1 the fringe system described above is displaced outwards from the central fringe and the fringes over a wide range of wavelength disappear as the load increases. However, fringes are again visible if a polarizing prism is placed in the beam of light entering M, and the prism turned so that the longitudinal axis of the specimen (and therefore the direction of the stress) is parallel or perpendicular to the direction of the electric vector.

These effects occur because the specimen becomes optically anisotropic under load with its axis of symmetry along the line of stress.

The birefringence induced by the stress results in the fringes observed in light polarized with its electric vector at right angles to the tension being displaced by a different amount from those observed in light polarized with its electric vector parallel to the tension. For a suitable stress, over a comparatively large wavelength range the minima of the fringes obtained in the one position of the polarizer correspond to the maxima obtained in the other position. Thus, in the absence of the polarizer, no fringes are visible. If the load is further increased the fringes again become visible, indicating that one system has continued to move relative to the other.

This is illustrated in the photograph, Figure 4 (Plate II), which was obtained by placing a Rochon double image polarizing prism in the collimated beam entering the interferometer, and a spherical lens in place of C for focusing the resultant two beams from M, polarized at right angles to each other, upon the spectrometer slit as two separate point images. A double fringe system is then produced. In the absence of any stress the two sets of fringes are in alignment. The fringes move away from the central maximum and move relatively to each other as the load is applied.

It is of interest to note that the spectrogram taken one minute after the removal of the load shows that the fringe systems have not yet returned completely to coincidence. This indicates some residual birefringence due to creep (Braybon 1949). That there was also a very small amount of birefringence present before the load was applied may be seen from close examination of the spectrogram taken prior to loading. This was due to the immediate past history of the specimen.

(ii) *Calculation of Stress Optical Coefficients*

Photoelastic properties of an amorphous solid are characterized by two stress optical coefficients B_1 and B_2, which at any given wavelength may be defined (Mueller 1935) by

$$\left.\begin{aligned}(\mu_x-\mu)&=\delta\mu_x=B_1P_z,\\(\mu_z-\mu)&=\delta\mu_z=B_2P_z,\end{aligned}\right\}\qquad\ldots\ldots(3)$$

where P_z is the pressure in the z direction.

μ is the refractive index of the unstressed medium and μ_x and μ_z are the resultant indices of refraction in the stressed substance for light with its electric vector perpendicular and parallel to the direction of pressure. The sign of P_z is thus negative for a tension. By subtraction:

$$B=B_1-B_2=(\mu_x-\mu_z)/P_z,\qquad\ldots\ldots(4)$$

where B is the relative stress optical coefficient. Most determinations are of the stress birefringence and merely give the constant B.

At any given wavelength, the values of $\delta\mu_x$ and $\delta\mu_z$ may be deduced, for differentiation of equation (1) keeping λ constant gives

$$\delta\mu=-(\lambda/2t)\,\delta n.\qquad\ldots\ldots(5)$$

The changes in fringe order, $\delta n_x, \delta n_z$, for light polarized with its electric vector respectively perpendicular and parallel to the stress P_z, are found by comparing (fringe-order, λ) curves obtained as above with those obtained in the unstressed condition. The corresponding refractive index changes are therefore:

$$\left.\begin{aligned}\delta\mu_x&=-(\lambda/2t)\delta n_x=B_1P_z,\\\delta\mu_z&=-(\lambda/2t)\delta n_z=B_2P_z,\end{aligned}\right\}\qquad\ldots\ldots(6)$$

and the resultant birefringence $\Delta\mu$ is given by

$$\Delta\mu=(\delta\mu_z-\delta\mu_x)=-(\delta n_z-\delta n_x)\lambda/2t=-BP_z.\qquad\ldots\ldots(7)$$

It should be noted that for a tension P_z is negative; for CR39 it was found that δn_x and δn_z are both positive and that $\delta n_x>\delta n_z$. Thus positive values of B_1, B_2 and B are obtained, indicating that, as with the majority of other substances showing stress birefringent properties, CR39 becomes positively birefringent under tension.

Furthermore, if the total load applied to the specimen is L, equations (6) and (7) may be rewritten, with $1/2tP_z$ replaced by $b/2L$, where b is the breadth of the specimen, a quantity which may be determined with less error than t, which is of the order of only a millimetre.

The application of a load will result in a change in the dimensions of the specimen, i.e. in b and t. For the low values of load employed, however, it was considered unnecessary to correct for this variation which for the load used is of the order of 0·05%.

§ 4. DISCUSSION

The method for the determination of stress optical coefficients bears some resemblance to that of Mach (1872) and of Twyman and Perry (1922). However, in the first of these two methods a Jamin interferometer was used and the method depended upon the use of a compensator block of the same material and upon a second measurement with specimen and compensator immersed in a fluid of known refractive index. In the second method a Michelson interferometer was used but the specimen was subjected to flexure and examined with monochromatic light.

The advantages of the present method are that only three sets of readings are required; measurements are made for small values of stress, reducing creep effects to a minimum; the specimen need not be very thin and fragile and its dimensions can be determined accurately. Estimation of the fringe shift for values of $\lambda \ll \lambda_0$ is somewhat difficult owing to the rapid rate of decrease of fringe order with decreasing wavelength, and results for the range may be somewhat less accurate. Obviously the method requires that the straight line through the points (λ_0, μ_0) and (λ_0', μ_0') be accurately determined and carefully drawn.

The method has been used successfully on CR39 over a wide range of wavelength. The work is being extended to other plastics and full details of the results will be published later.

ACKNOWLEDGMENTS

The author wishes to thank Professor A. M. Taylor for his advice and encouragement, and Mr. S. Weintroub for his assistance in the preparation of this paper.

REFERENCES

Braybon, J. E. H., 1949, *Nature Lond.*, **163,** 689.
Mach, E., 1872, *Ann. Phys., Lpz.*, Ser. II., **146,** 316.
Mueller, H., 1935, *Physics*, **6,** 179.
Pfund, A. H., 1941, *J. Opt. Soc. Amer.*, **31,** 679.
Twyman, F. and Perry, J. W., 1922, *Proc. Phys. Soc. Lond.*, **34,** 151.

Vibrations of Free Elliptical Plates

By MARY D. WALLER*

MS. received 16*th January* 1950

ABSTRACT. The systematic observations which have been made on the normal vibrations of free elliptical plates confirm Chladni's results. The nodal systems may be divided into the same four classes of symmetry as those of rectangular plates.

The correspondence which exists between the vibrations of elliptical, rectangular and circular plates is traced. Combined modes of vibration do not occur on elliptical plates as they occasionally do on rectangular plates.

§ 1. INTRODUCTION

THE only extensive experimental study of vibrating elliptical plates would appear to be that of Chladni. The 23 drawings given in *Die Akustik*, 1802, concern a plate whose axes are in the ratio of about 3/2 and the nodal systems are arranged systematically, as in the photographic records of the present paper. The text, §§ 142–157 includes data regarding the vibration frequencies of many other plates which may be converted from the notation of the chromatic scale by using the Zamminer table given in Rayleigh's *Theory of Sound*, I, 1894.

---

* Private address: 5 Gloucester Gate, Regent's Park, London N.W.1

Pavlík (1937) has used the Rayleigh–Ritz approximate method (Rayleigh 1894, 1911, Ritz 1909) to calculate the first seven natural frequencies of a rather wide stainless steel plate (ratio of axes 1·29) and to compare these with observation. The agreement is good for the lower frequencies, but for the sixth and last pair of frequencies the values given by theory exceed the observed values by 44% and 52% respectively. The theory, which involves laborious calculations, has not been tested on narrower ellipses.

It is evident that our information regarding the vibrations of elliptical plates must depend principally upon experiment, and it is desirable that photographs of the nodal designs produced should be placed on record.

The present work forms part of a more general study of the vibrations of plates which has not as yet reached completion. Details regarding the experimental technique have been given in previous publications (Waller 1949, and references therein).

The subject is approached through considerations of symmetry which make it easy to arrive at a general understanding of vibrations of plates of any given geometrical shape.

§ 2. SYMMETRY AND CLASSIFICATION OF THE NORMAL NODAL SYSTEMS

The elliptical plate possesses two mirror lines of symmetry, the major and the minor axis, and the rotational symmetry is 180°. Lines of symmetry divide the nodal pattern so that one half is the mirror image of the other. The fact that the lines of symmetry may be nodal or antinodal may be appreciated by examining the first nodal system of the first two rows of Figure 1 or 2 (see Plate). In the 1|1 system of the second row the lines are nodal and the displacements of corresponding points of the surface on either side of them are equal and out of phase. In the 0|2 system of the first row the lines are antinodal, that is to say they pass through the points of greatest displacement, and the displacements of corresponding points on either side of them are equal but in phase.

It is evident that all possible nodal systems can be divided into four classes according as (i) the centre is antinodal, (ii) the minor axis is nodal, (iii) the major axis is nodal, (iv) both the axes are nodal.

These classes correspond with those of rectangular plates where the two lines of symmetry are the two medians (Waller 1949).

All the nodal designs can be brought into self-coincidence by a rotation of 180° about the axis passing through the centre of, and perpendicular to the plane of, the plate.

§ 3. THE NORMAL NODAL SYSTEMS

Photographic records of the nodal designs obtained on a narrow plate (ratio of axes 2|1) and a wide plate (ratio of axes 5|4) are shown in Figures 1 and 2 respectively.

The arrangement of the nodal systems is the same as Chladni's and it corresponds also with that of rectangular plates (Waller 1949).

Each column, as indicated by the number above it, contains systems with a constant number of lines running in the direction of the shorter axis; and each row, as indicated by the number to the left of it, contains systems with a constant number of lines running in the direction of the longer axis. A nodal ellipse counts as two in the latter case. It will be seen that two nodal classes occur alternately in the first row and again in the third row, and the other two classes

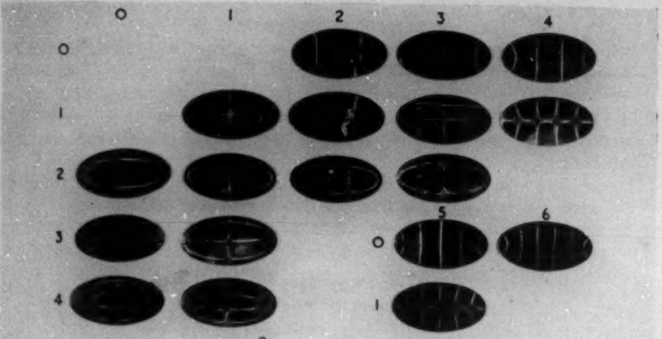

0
1
2
3
4
0
1
2
3
4
5
6
0
1

Figure 1. Ratio of axes 2/1

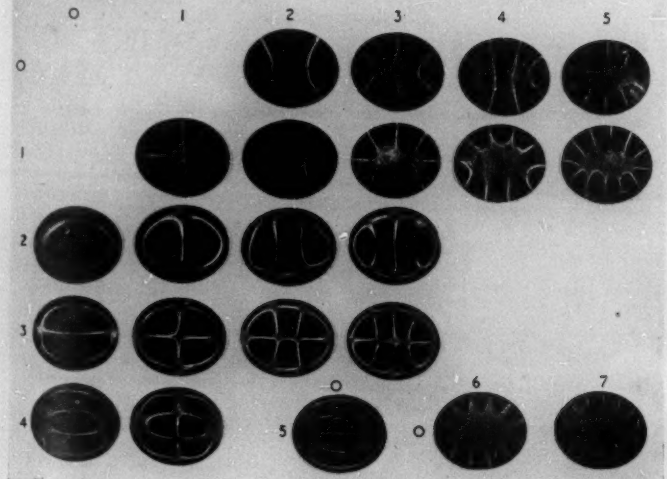

0
1
2
3
4
5
0
1
2
3
4
0
5
6
7
0

Figure 2. Ratio of axes 5/4

Normal nodal systems of free elliptical plates.

in the second row and again in the fourth row. Three spaces at the top left-hand corner must remain vacant. The diagram can be continued indefinitely downwards and to the right.

§ 4. THE NATURAL FREQUENCIES

Table 1, which is in two parts, has been arranged so as to correspond with the nodal systems of Figures 1 and 2 respectively. The frequencies are expressed in terms of the gravest tone taken as unity. Particulars regarding the dimensions and actual lowest frequencies in plates used are also included. The numbers in italics are approximate, having been estimated by extrapolation from curves derived from the earlier frequencies of any given row or column. That the frequencies are in general agreement with Chladni's approximate results can be seen by comparing Table 1 with Table 2, which has been compiled from his data concerning plates of the same shape as those used in the present experiments.

Table 1. Relative Frequencies of Normal Vibrating Modes of Free Elliptical Plates

2/1 Ellipse (see Figure 1)

	0	1	2	3	4	5	6
0			1*	2·58	4·7	*7·3*	*10*
1		1·77	3·27	5·68	8·29	11	
2	4·25	6·57	9·43	*12·6*			
3	10·6	*14*					
4	*17*	*22*					

* Actual frequency of brass plate, major axis 12·68 cm., minor axis 6·39 cm., thickness 1·62 mm., was 438 c/s.

5/4 Ellipse (see Figure 2)

	0	1	2	3	4	5	6	7
0			1*	2·45	4·28	6·66	9·39	*13*
1		1·07	2·59	4·34	*6·8*	9·6		
2	2·03	3·99	6·71	10				
3	4·42	7·41	10·7	*14*				
4	9·01	*12*						
5	*14*							

* Actual frequency of brass plate, major axis 12·60 cm., minor axis 10·18 cm., thickness 1·62 mm., was 414 c/s.

Table 2. Chladni's Vibration Frequencies (approximate)*

	2/1 Ellipse						5/4 Ellipse						
	0	1	2	3	4	5	0	1	2	3	4	5	6
0			1	2·5	4·7	7·4			1	2·5	4·5	7	9·2
1		1·6	3·3	5·6	8·3	11		1·1	2·5	4·5	7	9·2	
2	3·9	6·6	9·4	13	17	20	2	3·8	6·8	11	16	18	
3	10	14	18	20			4·8	7·4	11	14	19		
4	19	24					8·9	12					
5							14						

* Converted from chromatic scale notation.

§ 5. FURTHER REMARKS

(i) *Comparison with Circular Plate*

On a circular plate, which is the limiting shape of elliptical plates, nodal ellipses become circular and nodal hyperbolic lines become diameters.

The correspondence between elliptical and circular nodal systems may be traced on Figure 2 by proceeding as follows:

Study two rows at a time, the first with the second and the third with the fourth. Couple each design in the upper of the two rows with the one preceding it in the lower of the two rows. Then each pair of nodal systems will be found to correspond to a single nodal system on the circular plate (Waller 1938).

For example, 0|2 in the first row and 1|1 in the second row both correspond to the two-nodal diameter system of the circular plate, 0|3 and 1|2 to three diameters, and so on.

Similarly pairs of nodal systems can be picked out in the third and fourth rows, omitting the nodal ellipse which occupies the first place and which corresponds to the nodal circle. Thus the 2|1 and 3|0 systems both correspond with a nodal circle with one diameter, and so on.

The ratios of the frequencies of paired nodal systems can be studied in Table 1, noting how much nearer they are to unity for the wider than for the narrower plate, 1·07 for example as compared with 1·77 for the first pair. In the limiting case of the circle the ratio becomes equal to unity.

(ii) *Comparison with Rectangular Plates and with Bar*

As previously noted, the nodal systems of elliptical and rectangular plates are arranged in the same way. The nodal designs evidently become more nearly alike as the plates become narrower. On very narrow plates the only nodal systems met with in practice are those which can be placed in the first row of Figure 1.

These approximate to the nodal systems of free bars where the lines run at right-angles to the length and are divisible into two classes according as the centre is antinodal or nodal.

(iii) *Combined Modes of Vibration*

We may ask why among the photographic records of Figure 2 the 1|4 and 2|3 designs are so distorted. Turning to the second part, Table 1, it is found that the corresponding vibration frequencies are nearly equal, *6·8* and 6·71, respectively. The fact that the two modes have attempted to combine, somewhat imperfectly, indicates that the plate is not quite uniform and explains why the designs are so irregular.

Incidentally this provides by far the most sensitive means of detecting a want of uniformity in a plate. But unfortunately it is unlikely to be of any practical use since combination occurs so rarely and cannot be predicted.

We will conclude by mentioning one important difference between the nodal designs of rectangular and elliptical plates.

On rectangular plates a trace of a second vibrating mode can sometimes be detected in a given nodal design, which always belongs to the same class of symmetry and which does not impair the symmetry of the design. Such combined vibrations have not been noticed on elliptical plates. This problem of combined vibrations, due presumably to anticlastic coupling, is a difficult one. It can probably be studied most fruitfully by further experiments on plates of different shapes.

ACKNOWLEDGMENTS

I have again to thank the Council of the Royal Free Hospital School of Medicine and Dr. W. A. Leyshon for giving me the hospitality of the Physics Department and Mr. D. J. Morgan for kindly cutting the plates.

REFERENCES

Chladni, E. F. F., 1802, *Die Akustik*, 2nd unaltered edition 1830 (Leipzig: Breitkopf u. Härtel), §§ 142–157, figs. 179–201.

Pavlík, B., 1937, *Z. Phys.*, **107,** 458.

Rayleigh, Lord, 1894, *Theory of Sound*, I, 2nd edition (London: Macmillan), Chapter X; 1911, *Phil. Mag.*, **22,** 225.

Ritz, W., 1909, *Ann. Phys., Lpz.*, **28,** 737.

Waller, M. D., 1938, *Proc. Phys. Soc.*, **50,** 70; 1949, *Proc. Phys. Soc.* B, **62,** 277.

LETTERS TO THE EDITOR

Preparation of Alkali Metals in Glass

When studying the electrical conductivity of the alkali metals it is necessary, owing to their great chemical activity, to prepare specimens for experiment in sealed glass capillary moulds. It is usual to make these of common ' soft ' glass so that platinum electrodes (which do not react with the metals or become oxidized during glass-blowing) may be sealed in. In recent work it has been necessary to use two special glasses for the preparation of potassium and lithium specimens respectively; we wish here to report our experiences with them.

The electrical resistance of ' pure ' potassium specimens exhibited two distinct ' kinks ' in the low temperature region between 20° and 4° K. (MacDonald and Mendelssohn 1948, 1950). To determine whether this was an intrinsic property of the ideally pure metal or due in some way to a small impurity content several distillations were carried out yielding purer samples. A limit, however, as judged by the residual resistance at very low temperatures, was reached quite soon. It was presumed that this was due to impurity (sodium probably) being collected from the glass distillation container; in order to minimize this, ' Pyrex ' glass was used for making the container because of its lower sodium content. To effect further improvement a special glass (C 98) has been produced in which the only alkali ion present is potassium. This glass, of composition and properties shown in the accompanying Table, works well in a flame and can be drawn and blown easily. It is very resistant to devitrification (none appearing in five days at any temperature between 675° and 1,000° c.) and may be sealed directly to ordinary soda-lime-silicate glasses of about the same expansion coefficient. When this glass was used for the distillation vessel a useful improvement in residual resistance was immediately observed and, furthermore, the ' kinks ' were now found to have disappeared, suggesting very strongly that the resistance phenomenon was due fundamentally to small amounts of sodium impurity, although the fundamental physical source of the phenomenon is still obscure.

	SiO_2	B_2O_3	Al_2O_3	BaO	Na_2O	K_2O	Li_2O	Expansion coefficient (0–400° c.)
C 98	64	2	—	18	—	16	—	$9{\cdot}6 \times 10^{-6}$
C 51/65	69·2	—	—	18·2	—	—	12·6	$10{\cdot}3 \times 10^{-6}$
C 10	2·3	26·8	29·6	35·0	4·4	1·9	—	$9{\cdot}0 \times 10^{-6}$

Components are given in percentages by weight.

When attempting to work with lithium metal a serious obstacle is met with since the molten metal reacts very strongly with ordinary glass, forming almost instantaneously a black compound on the surface of the glass; the action generally proceeds so rapidly as to cause fracture of the glass itself. It has been suggested (see, for example, Justi 1948) that the high melting temperature of lithium is primarily responsible for this behaviour; since, however, sodium can readily be heated in glass to temperatures well above this value without appreciable reaction, this hypothesis seems untenable.

It was thought at first that the process might be one of replacement of the sodium in the glass by lithium, and, therefore, a special glass was made up which contained lithium as the only alkali ion. The well-known Lindemann glass is of this type (containing only lithium, beryllium and boric oxides), but it is difficult to work into the apparatus required by the experiments. The simple lithium-barium-silicate glass (C 51/65) of composition shown in the Table can be drawn into tubing (including capillary) by a skilled glass-worker, and although the glass has a low softening temperature and must be worked with care, the tube can be sealed directly to ordinary soda-lime-silicate glasses of similar expansion coefficient. Thin platinum electrodes can be sealed into the glass without cracking.

Unfortunately, tests soon showed that this glass was no more resistant to lithium than ordinary 'soft soda' glass. Recalling, then, that fused silica is also attacked in the same way by lithium, it was thought that the silica in the glass was probably the responsible constituent. An experiment was therefore made using tubing of the type employed in the construction of sodium discharge lamps; the tubing has an inner lining of a borate glass which is almost free from silica, and is known to be resistant to the attack of both sodium and potassium vapours. The composition of the glass (C 10) is also given in the Table. C 10 has proved sufficiently resistant to permit lithium to be cast successfully in it if the metal is solidified fairly rapidly (say 10 to 20 seconds) after running into the mould; thereafter a blackening does begin to occur. It is hoped in the future to make experiments with other glasses in which the silica content will be reduced to zero and the Na_2O and K_2O replaced by Li_2O.

Clarendon Laboratory, Oxford. D. K. C. MacDonald.

British Thomson-Houston Co. Ltd.,
Research Laboratory, Rugby. J. E. Stanworth.
4th April 1950.

Justi, E., 1948, *Ann. Phys., Lpz.*, **3**, 183.
MacDonald, D. K. C., and Mendelssohn, K., 1948, *Nature, Lond.*, **161**, 972; 1950, *Proc. Roy. Soc.* A (in the press).

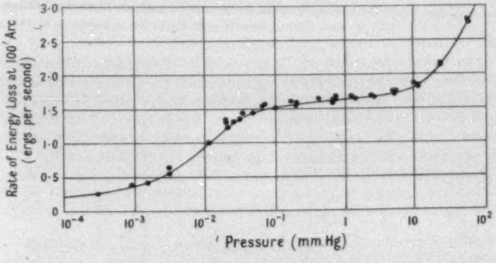
The Dissipation of Energy by a Pendulum Oscillating in Air at Low Pressures

The energy dissipation of a seconds pendulum has been measured by Atkinson (1938), who described experiments with different pendulums by means of which losses due to suspension, rod and bob were calculated. Such losses related to pendulums swinging in the open air, and these were found to be higher when the pendulums were inside cases. The loss of energy is due to four causes: (*a*) air resistance to the bob, (*b*) air resistance to the rod, (*c*) bending of the suspension spring, and (*d*) movement of the support. The combined loss due to causes (*c*) and (*d*) can be determined directly from measurements made with the pendulum in a vacuum.

The air pressure in the cases containing the 'Shortt clocks' is normally reduced to a value 30–20 mm. Hg, and it was decided to measure the losses of one of these pendulums at lower pressures and, by extrapolation to zero pressure, determine the losses due to the flexure of the suspension spring and the movement of the support.

The seconds pendulum of the standard Shortt clock seconds type, used for the tests, was carried by a spring ground from a solid strip of steel 2 mm. thick. The upper thick end of the strip was gripped between the two halves of a cylindrical trunnion while the lower thick end carried a steel pin which engaged with two hooks formed at the top end of the 8 mm. diameter invar pendulum rod. The bob, made of type metal, was a cylinder 9 cm. in diameter, 10 cm. long, weighing 6·5 kilograms. It was supported a little below its centre of gravity by means of an invar sleeve—pinned to the pendulum rod—carrying a brass collar of such a length that its upward expansion due to temperature compensated the corresponding downward expansion of the suspension spring and pendulum rod. The case was a copper tube of internal diameter 21·5 cm., fitted with heavy bronze flanges at the top and bottom carrying lugs by means of which it was solidly bolted to the wall. The trunnion, holding the top of the suspension, was supported by a four-legged bronze casting which stood on the inner edge of the top flange. This casting, which also carried the maintaining mechanism, was covered by a glass bell-jar, sealed to the outer edge of the flange. The base of the case was closed by a thick glass disc sealed to the bottom flange. When the pendulum was swinging, the underside of the cylindrical bob was 8·2 cm. from this plate, but the rod was extended downwards to carry a silver ' beat plate ' which moved to and fro just clear of the glass. This beat plate was engraved with a series of cross-lines spaced 1·67 mm. apart which was the appropriate interval to represent 5′ of arc. The movement of these lines, as the pendulum oscillated, was observed by means of a microscope. When illuminated by a suitably placed lamp, one of the plate lines came into view at the end of a swing, slowed to a standstill, and then swung ' out of focus ' again, followed by a similar movement of another line entering from the opposite side. It was quite a simple matter to read off these ' standstill positions ' on the eyepiece scale at any particular time or, alternatively, to observe the times when the lines ' stopped ' at particular scale divisions, or at the same division.

The rate of energy loss of a swinging pendulum is $I\omega\, d\omega/dt$, where I is the moment of inertia of the pendulum and ω the angular velocity when passing through its equilibrium position. The value of I for the pendulum used was $6{\cdot}852\times10^7$ G.C.S. units and, representing the total arc of swing by A, the rate of energy loss in ergs per sec. was $14{\cdot}28A\, dA/dt$. To determine the relation between the rate of energy loss and the gas pressure it was only necessary to measure dA/dt at 100′ total arc swing at various pressures.



The air pressure in the case having been adjusted to the required value, the arc of the pendulum was raised to 110′ or more by utilizing the normal impelling mechanism; the impulsing was then cut off. The times were recorded at which the total arc became 105′ and 95′, respectively, and dA/dt determined. The gas pressure was measured by means of an oil manometer at the higher pressures and by a McLeod gauge at the lower values. The results obtained are shown in the accompanying diagram.

They indicate that the energy loss due to the flexure of the particular spring used, together with the losses due to the movement of the suspension brackets, i.e. of the head of the case, is less than 0·27 erg per sec., and probably reaches a limiting value of 0·20 erg per sec.,

that the damping at 3×10^{-4} mm. Hg is only about one-sixth the damping at 10^{-1} mm. Hg, and that gas viscosity damping still persists at a pressure of 10^{-4} mm. Hg, which agrees with the fact that the quartz fibre vacuum gauge is operative at this pressure. The limiting value of 0·20 erg per sec., which represents the suspension losses, is considerably less than that previously estimated by the author—a value criticized by Atkinson as being too high—but the earlier figure was based on tests which were confined to pressures above 1 mm. Hg. The present tests also confirm the statement by Loomis, quoted by Atkinson, that the decrement of a half-second pendulum was reduced to one-half by decreasing the gas pressure from 0·025 mm. to 0·001 mm. Hg.

Department of Physics,
University College, Exeter.
5th April 1950.

W. H. SHORTT.

ATKINSON, E. C., 1938, *Proc. Phys. Soc.*, **50**, 721.

REVIEWS OF BOOKS

Higher Physics, by E. NIGHTINGALE. Pp xvi+808. (London: G. Bell and Sons Ltd., 1948.) 27*s*. 6*d*. Also in separate parts: Pt. I. *Mechanics and Properties of Matter*, 7*s*. 6*d*.; Pt. II. *Heat*, 7*s*. 6*d*.; Pt. III. *Light and Sound*, 10*s*.; Pt. IV. *Electricity*, 12*s*. 6*d*.

The clarity of style which has characterized Mr. Nightingale's many school texts in the past is in evidence in his latest effort. This book is of intermediate standard and is primarily intended for sixth form scholars.

The wide field of subject-matter which is covered within the confines of this single volume has necessitated considerable ' condensing ' in certain sections, but on the whole this process has been carried out quite efficiently. However, it is inevitable that some subdivisions become unduly compressed when room is found for such topics as the theory of flight, the quantum theory of spectra, etc.

An undesirable effect of the lag, so prevalent in the post-war era, between the preparation of an MS. and its publication is exemplified in the present instance by the author's failure to record the recent official acceptance (October 1946) of the absolute system of electrical units as the practical system, to the exclusion of the International Scale. Also no mention is made, rather surprisingly, of the M.K.S. system of units.

The author makes frequent use of the methods of dimensions in the various sections of the book and the use of graphs in exhibiting experimental results is also stressed in the text.

The errors noted are remarkably few, but one which should be noted is the mention, on page 633, of the use of an iron former for the suspended coil of a galvanometer.

Line diagrams which are capable of reproduction by the student are freely employed thoughout the text, and another feature which commends the book as one to be treasured by the scholarship candidate is the large selection of examination questions at the end of each chapter.

R. W. B. S.

Terrestrial Radio Waves: Theory of Propagation, by H. BREMMER. Pp. x+343. (New York, Amsterdam, London, Brussels: Elsevier Publishing Co., Inc.; British Agents: Cleaver-Hume Press, 1949).

The author asks us in his Preface to bear in mind that his aim is to describe the mathematical-physical methods for the computation of transmitter fields. The statement of the problem is presented briefly in the opening pages and looks deceptively simple: it is required to compute the radiation field of a Hertzian dipole above a spherical earth. The influence of the ionosphere and the refracting lower atmosphere is deferred until Part II of the book. It takes some hundred odd pages of concisely written mathematical analysis to arrive at the final formulae for the first part of the problem in a form suitable for numerical work, and the reader will readily agree with Nicholson that the problem was one of the most difficult at one time facing the theorist. The author and his colleague Professor van der Pol were foremost in reducing its solution to a practical form for numerical work ; their original research papers, published in the *Philosophical Magazine* in 1937–1939, have long been recognized as the " textbook " of workers in this field. Indeed, much of Part I will be familiar to those who have read these papers.

The apparent ease with which the author systematically takes us through pages of complicated mathematical analysis could only come after long experience and familiarity with his subject; his emphasis on the physical interpretation at each stage is an outstanding example of the presentation of mathematical physics. Even so, it is to the experienced theoretician that his descriptive passages are directed, and it is unlikely that the average radio engineer has the necessary mathematical equipment for their appreciation. Nevertheless the debt owed by the radio engineer, as shown by his extensive use of the results of this work in the past, was aptly expressed by T. L. Eckersley in his concluding remarks to the Radiolocation Convention in 1946 when he said that " there is nothing so practical as a good theory ".

Without much preliminary, a beginning is made to the solution of the point source and spherical earth problem. Starting with the classical solution, in the form of a series of zonal harmonics, the author takes us through the various transformations and approximations with systematic thoroughness, eventually leading to the familiar residue series. An instructive account of the theories of steepest descent and stationary phase is of general interest in the relationship of ray to wave theory, and is more clearly discussed here than in most books on the subject. The author does well in giving prominence to the physical significance of geometric-optical approximations and of Watson's transformation, which might easily have remained as so much mathematics.

The concluding chapter of Part I will be particularly welcomed by the radio engineer whose interest is in the calculation of field strengths. The specific examples chosen and worked through in detail will be of assistance in using the formulae developed earlier in the book and now collected together in a convenient form for computation. Curves covering a wide range of frequencies are also included.

Part II of the book is devoted to a discussion of the influence of the ionosphere and refraction by the troposphere. With the simplifying assumption of a stratified atmosphere whose refractive index varies only with height, a comprehensive discussion is given of both the residue series and geometric-optical solutions. The reader will find an enlightening account in Chapter VII of the method by which the spherical earth problem can be replaced by one with a " flat " earth and a modified index of refraction. Closely allied to this is Eckersley's method of taking refraction into account to a first approximation by using a " fictitious " radius of earth which is usually larger than the actual radius. During the war years this theory has been supplemented by the more refined mode theory of super-refraction which explains the abnormally long ranges experienced by microwave radar during favourable meteorological conditions. The complexities of the eigenvalue problem involved in the evaluation of height functions receives detailed attention ; the discussion of the W.K.B. and the more accurate Krammer phase integral approximation is of general interest to the mathematical physicist.

This brief outline of the book cannot possibly do justification to its wealth of detail. It is without doubt a book for the propagation specialist who requires a fundamental understanding of his subject. There is, however, much of general interest for the theoretician.

W. WALKINSHAW.

Oscillations of the Earth's Atmosphere, by M. V. Wilkes. Pp. ix+74. Cambridge Monographs on Physics. (Cambridge : University Press, 1949). 12*s*. 6*d*. net.

This excellent little volume by the Director of the University Mathematical Laboratory, Cambridge, deals with the oscillations produced daily in the atmosphere by the sun and the moon, as revealed mainly by the solar and lunar daily variations of barometric pressure. The subject may be said to have started with Torricelli's invention of the barometer and Newton's gravitational theory of the tides, greatly developed by Laplace, who in his last years devoted much attention to atmospheric oscillations. Kelvin showed the importance of resonance in greatly magnifying the semidiurnal atmospheric oscillations due to the sun's thermal and tidal action, so that the lunar tide in the barometer, unlike that in the sea, is relatively inconspicuous. The present data on the subject are described, and also Laplace's tidal theory. The last and most novel part of the book is Chapter V which gives the modern theory of these oscillations and their resonance, a theory begun by Taylor and Peteris and since then much developed by the author and Dr. K. Weekes. The book should interest many physicists, mathematicians and meteorologists and stimulate further developments of which the subject still stands in much need. S. C.

Sound (*A University Text-book of Physics, Volume II*), by J. H. Poynting and J. J. Thomson, revised by W. S. Tucker. Pp. viii + 251. Tenth Edition. (London: Charles Griffin, 1949.) 20*s*.

During the last 50 years nine editions of this well-known textbook have been published. It is not surprising that in this present edition the reviser has found it necessary to write effectively a new book. The text has been rearranged and the theory simplified by the use of calculus methods. The reviser claims that the book has been written to remedy a deficiency in present University courses, where ' Sound ' " is frequently relegated to the end, and is largely crowded out ", and has been planned for students taking honours courses in branches other than physics, viz. those who have chosen physics as a subsidiary subject for their degrees.

The text mostly divides into two parts—pure and applied acoustics. The former follows normal procedure and requires little comment except that an improvement could be made in the definition of terms introduced and in the physical–mathematical analysis relation. In the applied acoustics part the question of relevancy of subject matter is much a question of opinion but it would appear that the text would gain in clarity and cohesion by considerable pruning and rearrangement. A considerable proportion of the book, about 34 pages, is devoted to the hot wire microphone and its application to sound location of aircraft and guns, and a further 10 pages to the acoustical design of aircraft sound locators. A discussion of the siren takes thirteen pages. Although aircraft noise and its measurement is considered in eight and a half pages there is little if anything on other noises and their measurement and suppression. Certain subjects are introduced in an unusual way, e.g. the Doppler effect through aircraft sounds without further mention of application, and the Whispering Gallery effect through observations on large sound reflectors for the detection of approaching aircraft. Some of the treatment appears to be unnecessarily dated, e.g. the recorded frequency spectrum of gramophone recordings is quoted as 50–5,000 c/s. and the last reference given is 1926. No mention is made of the recent improvements in wire and tape magnetic recording. Readers may argue however that such details are unnecessary in a textbook of this standard since they do not add to the fundamental concepts.

An erroneous impression may be given by certain statements such as : on p. 93, the uses of the tuning fork as a present standard of frequency for radio and clock control; on p. 242, the highly absorbent walls of broadcast studios; and on p. 134, the description of the mechanics of the electronic musical instruments. There seems to be some misunderstanding on the relative positions of the sound and vision gates on film projectors.

On musical questions, scale, consonance, etc., tend to be considered as a result of numerical arrangement. Numerical approach has interest but is mostly sterile.

The last chapter in the book, viz. " The Acoustics of the Concert Hall " could be better titled. Only a quarter of the printed matter refers to this question.

In a future edition an improved text could be obtained by a rearrangement of the material, with the deletion or considerable reduction of certain sections and the incorporation of further diagrams and photographs on the pure acoustical side. The addition of some thought-provoking problems at the end of chapters is also suggested. H. D. P.

PUBLICATION RECEIVED

High Vacuum Technology, by A. S. D. Barrett.

Abridged version of the 23rd Anniversary Address given to the King's College Engineering Society at the Institution of Mechanical Engineers on 10th March 1949. (London : W. Edwards & Co. Ltd.)

CONTENTS FOR SECTION A

PAGE

Dr. M. W. **Feast**. The Schumann–Runge O_2 Emission Bands in the Region 3100 A.–2500 A. . . . 549

Dr. M. W. **Feast**. New O_2^+ Second Negative Bands: A Note on O_3 and O II Emission Spectra . . . 557

Dr. M. W. **Feast**. The Spectra Emitted by the High Voltage Arc in Nitrogen, Hydrogen, Nitrogen–Hydrogen Mixtures and Ammonia . . . 563

Dr. M. W. **Feast**. Rotational Analysis of the (1, 0) Band of the N_2 First Positive System . . . 568

Dr. H. D. **Evans**. An Absorption Comparison of the β-Particle Spectra of $^{207}_{81}$AcC″ (allowed), $^{210}_{83}$RaE (second forbidden) and 3·5 yr.-$^{204}_{81}$Tl (third forbidden) . . . 575

Mr. G. R. **Baldock**. Excited Electronic Levels in Conjugated Molecules—V: A Valence Bond Estimation of Energy Levels in Aromatic Hydrocarbon Molecules . . . 585

Mr. J. W. **Cook**, Miss R. **Schoental** and Mr. E. J. Y. **Scott**. Relation between Bond Structure and the Longest Ultra-violet Absorption Band of Polycyclic Aromatic Hydrocarbons . . . 592

Mr. S. T. **Butler**. The Scattering of High Energy Charged Particles by Thin Foils of Matter . . . 599

Dr. M. G. **Nooh** and Mr. S. R. **Haddara**. Penetrating Showers at High Altitude. 606

Dr. J. H. **van der Merwe**. On the Stresses and Energies associated with Inter-Crystalline Boundaries . . . 616

Mr. R. B. **Dingle**. The Theory of the Propagation of First and Second Sound in Helium II.—Energy Theorems and Irreversible Processes . . . 638

Mr. J. D. **Lawson**. The Angular Distribution of Synchrotron Target Radiation: A Preliminary Experimental Study . . . 653

Letters to the Editor:

- Mr. D. J. **Lees** and Mr. L. H. **Metcalfe**. Measurement of Polar Diagram of Synchrotron Gamma Radiation . . . 661
- Mr. F. K. **Goward** and Mr. J. J. **Wilkins**. Identification of Photo-Disintegration Stars in Nuclear Emulsions . . . 662
- Mr. J. J. **Wilkins** and Mr. F. K. **Goward**. Identification of Nitrogen Photo-Disintegration Stars in Nuclear Emulsions . . . 663
- Dr. A. P. **French** and Dr. P. B. **Treacy**. Alpha-Particles from ^{27}Al + D . 665
- Dr. A. P. **French**, Dr. P. **Meyer** and Dr. P. B. **Treacy**. Alpha-Particles from ^{19}F Bombarded with Deuterons . . . 666
- Dr. Joan M. **Freeman**. The Energy-Release in some (pα) Reactions in Light Nuclei . . . 668
- Prof. E. A. **Stewardson** and Mr. H. F. **Zandy**. The L–M and K–M Discrepancies in the Rare Earths . . . 670
- Mr. R. W. K. **Honeycombe**. Inhomogeneity of Deformation in Metal Single Crystals . . . 672
- Dr. M. **Afaf**. Singlet System B of ZrO . . . 674

Reviews of Books . . . 675

Contents for Section B . . . 678

Abstracts for Section B . . . 679

ABSTRACTS FOR SECTION A

The Schumann–Runge O_2 Emission Bands in the Region 3100 A.–2500 A., by M. W. FEAST.

ABSTRACT. Rotational and vibrational analyses of the emission Schumann–Runge O_2 bands in the region 3100 A. to 2500 A. are presented. Tables of wave-numbers are given for the (1, 12), (0, 11), (1, 11), (0, 10), (1, 10), (1, 9), (2, 9), (1, 8), (2, 8), (2, 7) bands. The origins of the bands, the rotational differences, and constants for the levels $v''=7$ to 11 are given. Three bands, found in a high voltage arc in oxygen between platinum electrodes, are attributed to PtO.

New O_2^+ Second Negative Bands: A Note on O_3 and O II *Emission Spectra*, by M. W. FEAST.

ABSTRACT. The spectrum of an electrodeless discharge in pure oxygen has been studied. New bands of the O_2^+ Second Negative system have been found and the intensity distribution in this system is discussed; subsidiary parabolae as well as the main Franck-Condon parabola are obtained as is predicted by the theory. Emission bands reported by Johnson as due to O_3 are due to O_2^+ and O II. Some abnormal intensities of O II lines excited in the electrodeless discharge are noted. The electrostatic and electromagnetic types of excitation are briefly compared.

The Spectra Emitted by the High Voltage Arc in Nitrogen, Hydrogen, Nitrogen-Hydrogen Mixtures and Ammonia, by M. W. FEAST.

ABSTRACT. The spectra emitted by the high voltage arc in nitrogen, hydrogen, hydrogen–nitrogen mixtures and ammonia have been studied.

The appearance of the N_2 Second Positive bands at various nitrogen gas pressures is discussed and the divergence of rotational and vibrational temperatures is noted. Traces of oxygen greatly reduce the strength of the N_2^+ bands emitted by the arc in nitrogen. In N_2–H_2 mixtures the NH 3360 A. band is found to be emitted strongly only when hydrogen is present at a very low partial pressure. An attempt is made to explain the various observations.

Rotational Analysis of the (1, 0) *Band N_2 First Positive System*, by M. W. FEAST.

ABSTRACT. The (1, 0) band of the N_2 First Positive system has been photographed at a dispersion of about 1·2 cm^{-1}/mm. The wave numbers and quantum numbers of the lines, and the rotational term differences in the upper and lower states are given. The spin tripling in the $^3\Sigma$ state and the Λ doubling in the $^3\Pi$ state are evaluated.

An Absorption Comparison of the β-Particle Spectra of $^{207}_{81}AcC''$ (allowed), $^{210}_{83}RaE$ (second forbidden) and 3·5 *yr.-$^{204}_{81}Tl$ (third forbidden)*, by H. D. EVANS.

ABSTRACT. Experiments have been described in which a direct comparison is made by an absorption method of the β-particle spectra of actinium C″, radium E and 3·5 yr. thallium representing allowed, second forbidden and third forbidden transitions respectively.

New values of 623 ± 4 mg/cm^2 and 300 ± 3 mg/cm^2 are proposed for the absorption limits in aluminium of the β particles emitted by actinium C″ and thallium and it is shown that, of the three elements, radium E emits most and actinium C″ least low energy electrons.

An attempt has been made to derive energy distributions from the absorption curves, but the results for radium E depart markedly from those obtained with magnetic spectrographs. The general findings, however, agree with those of other workers who have attempted the same procedure.

The formula used in the derivation of the energy distribution includes an approximation which necessarily leads to an apparent deficiency of low energy particles, but the accuracy of the data is not sufficient to permit the use of a better approximation. Strictly for purposes of comparison, however, the method has obvious applications.

Excited Electronic Levels in Conjugated Molecules—V: A Valence Bond Estimation of Energy Levels in Aromatic Hydrocarbon Molecules, by G. R. Baldock.

ABSTRACT. This note is concerned with the calculation of the energy of the ground state and the lowest excited state for hydrocarbons of the condensed ring type. The difference between these two energies is the energy of the first electronic transition. The normal valence bond method is used, taking only the Kekulé structures into consideration. The results are compared with the experimental absorption wavelengths, and are in good agreement, giving a value of the exchange integral α between $-1{\cdot}8$ and -2 electron volts.

Relation between Bond Structure and the Longest Ultra-violet Absorption Band of Polycyclic Aromatic Hydrocarbons, by J. W. Cook, R. Schoental and E. J. Y. Scott.

ABSTRACT. A comparison has been made between the values expressed in wave numbers of the longest ultra-violet absorption bands of a series of polycyclic aromatic hydrocarbons and the minimum number of quinonoid rings contained in their molecules. It is found that the hydrocarbons fall into series of groups related to this number of quinonoid rings.

The Scattering of High Energy Charged Particles by Thin Foils of Matter, by S. T. Butler.

ABSTRACT. The distribution of high energy charged particles scattered by a thin foil has been investigated and compared with the formula for single scattering. It is shown that at sufficiently large angles the distribution approaches that for single scattering and can be expressed in the form of a rapidly convergent series of which the first term is identical with the single scattering formula. The remaining terms of the series form an accurate expression for the deviation from true single scattering. The simplified case of projected scattering has also been considered.

Penetrating Showers at High Altitude, by M. G. Nooh and S. R. Haddara.

ABSTRACT. A study has been made of cloud chamber photographs of penetrating cosmic-ray showers, originating in lead placed over the cloud chamber and in a lead plate within the cloud chamber, obtained at altitude 3,572 m. in a magnetic field of about 1,000 gauss.

The use of well-separated successive lead layers allowed the association of penetrating events to be demonstrated, and there is some evidence of a general reduction of energy involved in events in the lower lead layers as compared with those in the top layer. Mesons, protons and particles of higher charge were identified in the penetrating showers, and a mean range for non-coulomb scattering of shower particles is given. Evidence is also given of the apparent starting and stopping of charged particles in the lead plate at the centre of the cloud chamber.

On the Stresses and Energies associated with Inter-Crystalline Boundaries, by J. H. van der Merwe.

ABSTRACT. Models, largely base don the assumptions introduced by Peierls and Nabarro in dealing with a single dislocation, are used in calculations on three types of intercrystalline boundaries, namely, (I) a boundary due to a difference of atomic spacing, (II) a twist boundary, and (III) a symmetrical tilt boundary. With these models the resolution of the boundary into a sequence of dislocations is a natural consequence of the analysis, which also yields the expressions for the stresses, atomic displacements and energies associated with the boundary, as functions of the angle of tilt, etc. By allowing the distance between dislocations to tend to infinity, these expressions reduce to the corresponding ones for single dislocations. The outstanding feature of the interfacial energy is that it increases initially very rapidly with the angle of tilt, etc. An application of the results to the theory of oriented overgrowths, developed by Frank and the present author, is described. The validity of the assumptions and approximations involved and the advantages of the treatment as compared with those of other writers are discussed.

The Theory of the Propagation of First and Second Sound in Helium II.—Energy Theorems and Irreversible Processes, by R. B. Dingle.

ABSTRACT. The first part of this paper contains a discussion of the conditions determining whether sound waves in helium II are propagated isothermally or adiabatically. In the remainder of this part an expression is found for the energy flux from a closed region, and the impedance concept is applied to waves of second sound.

The second part of the paper is devoted to the study of the effects of irreversible processes on the propagation of waves of first and second sound. The irreversible processes discussed are those due to viscosity and thermal conduction, both in an extended fluid and in a narrow tube.

The Angular Distribution of Synchrotron Target Radiation: A Preliminary Experimental Study, by J. D. Lawson.

ABSTRACT. The angular distribution of radiation from a synchrotron has been measured with an ionization chamber for a number of different target materials and thicknesses, and the results have been compared with a simple theory in which the radiation from a parallel beam of electrons striking a thin plate has been calculated. Detailed comparison with theory is not possible, because the precise way in which the electrons hit the target cannot be determined, nevertheless qualitative explanations have been found for most of the effects observed. The most important of these is the relative insensitivity of forward intensity and beam width to target size and material.